AL/HR-TR-1993-0131

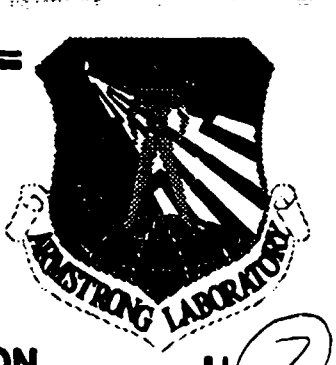


IMAGE QUALITY AND THE DISPLAY MODULATION TRANSFER FUNCTION: EXPERIMENTAL FINDINGS

Ronald J. Evans

University of Dayton Research Institute
300 College Park Avenue
Dayton, OH 45469

HUMAN RESOURCES DIRECTORATE
AIRCREW TRAINING RESEARCH DIVISION
6001 S. Power Road, Bldg 558
Mesa, AZ 85206-0904

September 1993

Final Technical Report for Period January 1991 - March 1992

Approved for public release; distribution is unlimited.

AIR FORCE MATERIEL COMMAND
BROOKS AIR FORCE BASE, TEXAS

ARMSTRONG LABORATORY

NOTICES

This technical report is published as received and has not been edited by the technical editing staff of the Armstrong Laboratory.

When Government drawings, specifications, or other data are used for any purpose other than in connection with a definitely Government-related procurement, the United States Government incurs no responsibility or any obligation whatsoever. The fact that the Government may have formulated or in any way supplied the said drawings, specifications, or other data, is not to be regarded by implication, or otherwise in any manner construed, as licensing the holder, or any other person or corporation; or as conveying any rights or permission to manufacture, use, or sell any patented invention that may in any way be related thereto.

The Office of Public Affairs has reviewed this report, and it is releasable to the National Technical Information Service, where it will be available to the general public, including foreign nationals.

This report has been reviewed and is approved for publication.

BYRON J. PIERCE
Project Scientist

DEE H. ANDREWS
Technical Director

LYNN A. CARROLL, Colonel, USAF
Chief, Aircrew Training Research Division

REPORT DOCUMENTATION PAGE

Form Approved
OMB No. 0704-0188

1. AGENCY USE ONLY (Leave blank)	2. REPORT DATE	3. REPORT TYPE AND DATES COVERED
	September 1993	Final January 1991 - March 1992

4. TITLE AND SUBTITLE

Image Quality and the Display Modulation Transfer Function: Experimental Findings

5. FUNDING NUMBERS

C - F33615-90-C-0005
PE - 62205F
PR - 1123
TA - 03
WU - 85

6. AUTHOR(S)

Ronald J. Evans

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)

University of Dayton Research Institute
300 College Park Avenue
Dayton, OH 45469

8. PERFORMING ORGANIZATION REPORT NUMBER

9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)

Armstrong Laboratory (AFMC)
Human Resources Directorate
Aircrew Training Research Division
6001 S. Power Road, Bldg 558
Mesa, AZ 85206-0904

10. SPONSORING/MONITORING AGENCY REPORT NUMBER

AL/HR-TR-1993-0131

11. SUPPLEMENTARY NOTES

Armstrong Laboratory Technical Monitor: Dr. Byron J. Pierce, (602) 988-6561.

12a. DISTRIBUTION/AVAILABILITY STATEMENT

Approved for public release; distribution is unlimited.

12b. DISTRIBUTION CODE

13. ABSTRACT *(Maximum 200 words)*

Image quality metrics represent an attempt to quantify differences in the quality of the transmission and display of visual information. This report focuses on components in the image transfer process which contribute to image quality as well as tasks through which image quality may be empirically defined. Components consist of the content of the original image, display device characteristics, and observer characteristics. Special attention within these three components is given to the display Modulation Transfer Function (MTF) which has traditionally been the major contributor to image quality metrics. Ambiguities exist in the definition and measurement of display MTFs and these problems are discussed as they pertain to image quality. Additional discussion includes the use of threshold versus suprathreshold tasks as empirical measures of image quality and the use of the Contract Sensitivity Function (CSF) versus the MTF of the eye in image quality metrics. An argument is presented which questions the use of either the CSF or MTF for suprathreshold tasks. In order to test the use of display MTFs in metrics, a methodology is described for digitally filtering images with filter representing hypothetical display MTFs. Although this method permits a subset of display MTFs to be compared, further efforts are required to compare MTFs which exhibit a crossover effect in the spatial frequency domain. Finally, empirical observations suggest that other display parameters (e.g., luminance) must be weighted more heavily in image quality metrics. A factorial approach to manipulate both the MTF and the display luminance is suggested to study this problem.

14. SUBJECT TERMS			15. NUMBER OF PAGES
Contrast sensitivity function	Fourier transform	Modulation depth	92
Convolution filter	Image quality	Modulation transfer function	16. PRICE CODE
Display quality	Image quality metric	Spatial frequency	

17. SECURITY CLASSIFICATION OF REPORT	18. SECURITY CLASSIFICATION OF THIS PAGE	19. SECURITY CLASSIFICATION OF ABSTRACT	20. LIMITATION OF ABSTRACT
Unclassified	Unclassified	Unclassified	UL

CONTENTS

LIST OF FIGURES

LIST OF FIGURES (Concluded)

PREFACE

The present effort was conducted in support of the Armstrong Laboratory/Aircrew Training Research Division (AL/HRA) research concerning image quality in simulator displays. The goal of this effort was to further the effort in developing a quantitative model to be used in predicting the quality of simulator display systems. The project was conducted under Work Unit 1123-03-85, Flying Training Research Support. Research support was provided by the University of Dayton Research Institute under Contract No. F33615-90-C-0005. The contract monitor was Ms. Patricia Spears, AL/HRAP.

The goal of this specific research effort was (a) identify critical components within the image display process which would be required in an image quality metric, (b) critically examine the measurement and use of the display modulation transfer function in currently used metrics, and (c) identify a methodology for experimentally studying image quality from a multidimensional perspective.

The author wishes to express thanks to Ms. Marge Keslin for final edit of the manuscript.

IMAGE QUALITY AND THE DISPLAY MODULATION TRANSFER FUNCTION: EXPERIMENTAL FINDINGS

INTRODUCTION

Modern simulator displays are built using a variety of technologies, making the choice of display type for a particular simulator a complex process. The criteria used in choosing a display system for a particular task must include cost, physical size limitations, maintainability, and user satisfaction or acceptability. Cost and physical size factors can be specified objectively while estimates of maintainability are somewhat vague, requiring tasks such as a reliability analysis and estimates of personnel required to maintain specific functions. It is user satisfaction/acceptability, based upon the quality of the imagery displayed, which eludes comprehensive measurement or analysis of any generality.

Acceptability of imagery/display systems, whether it be for specifically trained tasks or aesthetic appeal, is studied within the realm of image quality research. Image quality research focuses on the prediction, explanation, and understanding of physical and psychological factors which determine the quality of imagery.

One primary goal in image quality research is to construct scalar measures or metrics which will order imagery along a single dimension, a dimension denoting the "goodness" or quality of the image. The measures or metrics will be analytic derivations based upon a composition of factors from the image, the display device, and the observer.

Figure 1 emphasizes the sequential nature of these subsystems (the original image, the display device, and the observer) in their contribution to the overall imag. quality process. In Figure 1, these subsystems are shown to be defined under a specific task or set of tasks under the assumption that perceived image quality is a task dependent quantity.

The use of the term "goodness" of an image brings to mind one of the major shortcomings noted by most researchers in image quality--that of objectively or operationally defining image quality. Two distinct measures have dominated the empirical work, viewer preference and psychophysical performance. Viewer preference typically involves ratings or ranking of the aesthetic value of imagery (e.g., Kusaka, 1989; Zetsche & Hauske, 1989). Performance tasks, consisting of reaction time and accuracy measures, require observers to make discriminations under degraded viewing conditions (e.g., Task, 1979). The relationship between these two empirical measures has received little theoretical or

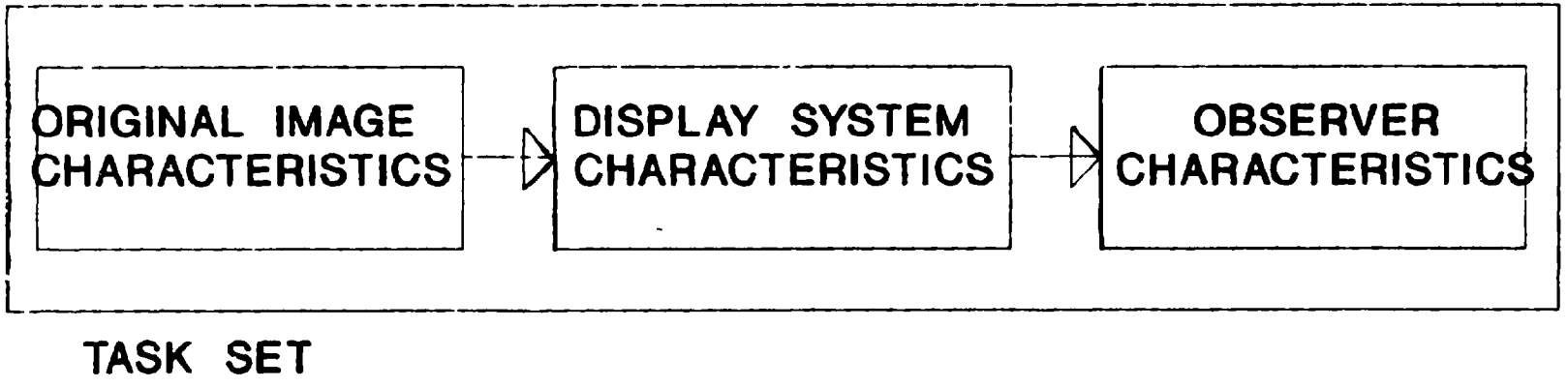


Figure 1
Systems Approach to the Study of Image Quality

empirical discussion. In this report, the relationship between the two measures will be discussed only indirectly, based upon threshold and suprathreshold stimulation of the human visual system.

A second major division can be made delineating imagery from the reading of text (e.g., Roufs & Bouma, 1980). The relationship between the quality of these two types of imagery is also unclear, text quality inherently being a serial, task-driven scenario dominated by foveal performance, and imagery belonging to parallel, spatially based processes including both foveal and peripheral vision.

The experimental work under consideration in this report includes only picture-type imagery. In addition, observer preference, typically associated with suprathreshold behavior, is used as the empirical measure of image quality and it will be assumed to have face validity. It is evident, however, that there is a need to clarify the relationship between the observer preference-based definition of image quality and the performance-based definition, typically associated with threshold behavior. In this report, this topic is discussed more thoroughly in the section describing the Contrast Sensitivity Function (CSF) and the Modulation Transfer Function (MTF) of the eye.

In flight simulation technology, advances have provided a variety of display devices from which to choose. These devices range from helmet-mounted displays to large dome displays. Display parameters vary dramatically across this range of devices. For example, a helmet-mounted display physically subtends less than a square foot but provides high brightness capability (e.g., 30 fL). The surface area of a large dome display may extend over a thousand square feet and the luminance may be on the order of a moonlit night (< 1/2 fL). In addition, the physical viewing distance for each of these displays can have quite disparate effects on the human visual system. This includes both static properties, such as the focusing of the eyes, and dynamic properties such as optical flow rates.

The multidimensional characteristics of display devices can only be critiqued with the knowledge of how the visual system employs or combines these parameters. Research into understanding how physical display parameters interact with the human visual perception is fragmented, concentrating on individual components of the process (e.g., contrast, luminance, and color thresholds). This analytical approach, however, is practical only because of the complexity of the human visual system and brain as well as the combinatorial ranges of the individual factors within the image/display which affect perception.

At the receiving end of the visual communications channel, the human visual system affords the highest rate of information transfer between machine and human. Because of this fact and the advances in the ability to store and transmit information visually, there is a renewed interest in developing numerical measures or metrics which capture the perceived quality of this visual transfer of information.

The ultimate goal of the metric approach would be to provide a unidimensional scale of image quality where display devices could be compared with one another for universal use. More practically, though, it is likely that metrics could be developed for a variety of visual tasks (e.g., monitoring, dynamic tracking, target detection, viewer preference). The image quality from a display device would then be a point in a multidimensional space with the dimensionality governed by the number of tasks.

Experimental and operational definitions of image quality rely upon human performance data and observer preference or ratings for empirical validation (e.g., Snyder, 1985). As mentioned, though, from an analytical viewpoint, the aesthetics of an image may have little to do with an observer's ability to detect or react to a target presented within an image. Experimentalists should be aware of this potential delineation in the two types of data (performance vs. preference) and generalize appropriately when collecting data and testing metrics.

To date, much of the experimental work performed in image quality uses correlational techniques to affirm the validity of newly developed metrics. Because of the types of data being used (e.g., preference data which may have only ordinal validity), use of interval-level statistics such as a Pearson Correlation coefficient are tenuous at best.

Currently, it is fairly straightforward to form a "partial ordering" of image quality (i.e., this display is at least as good as that display) based upon the comparison of a single physical display dimension (e.g., resolution, luminance) given all other display parameters are equal across displays. The more pertinent questions in image quality, however, are in evaluating trade-offs across display dimensions (e.g., luminance versus resolution) when determining image quality. Figure 2 is an example of trade-offs across dimensions. In Figure 2, the Michelson Contrast or modulation depth (introduced in the subsection "Luminance Modulation: Contrast Measures") has been measured for two display devices, the Display for Advanced Research Training (DART) (Thomas, Reining, & Kelly, 1990) and a limited field-of-view (LFOV) dome display in use at the Armstrong Laboratory, Williams AFB, AZ. If either display had better modulation at all spatial frequencies, that display device would be assumed to provide better image quality given all other parameters were equal (e.g., luminance, color, temporal properties). However, in Figure 2, the DART

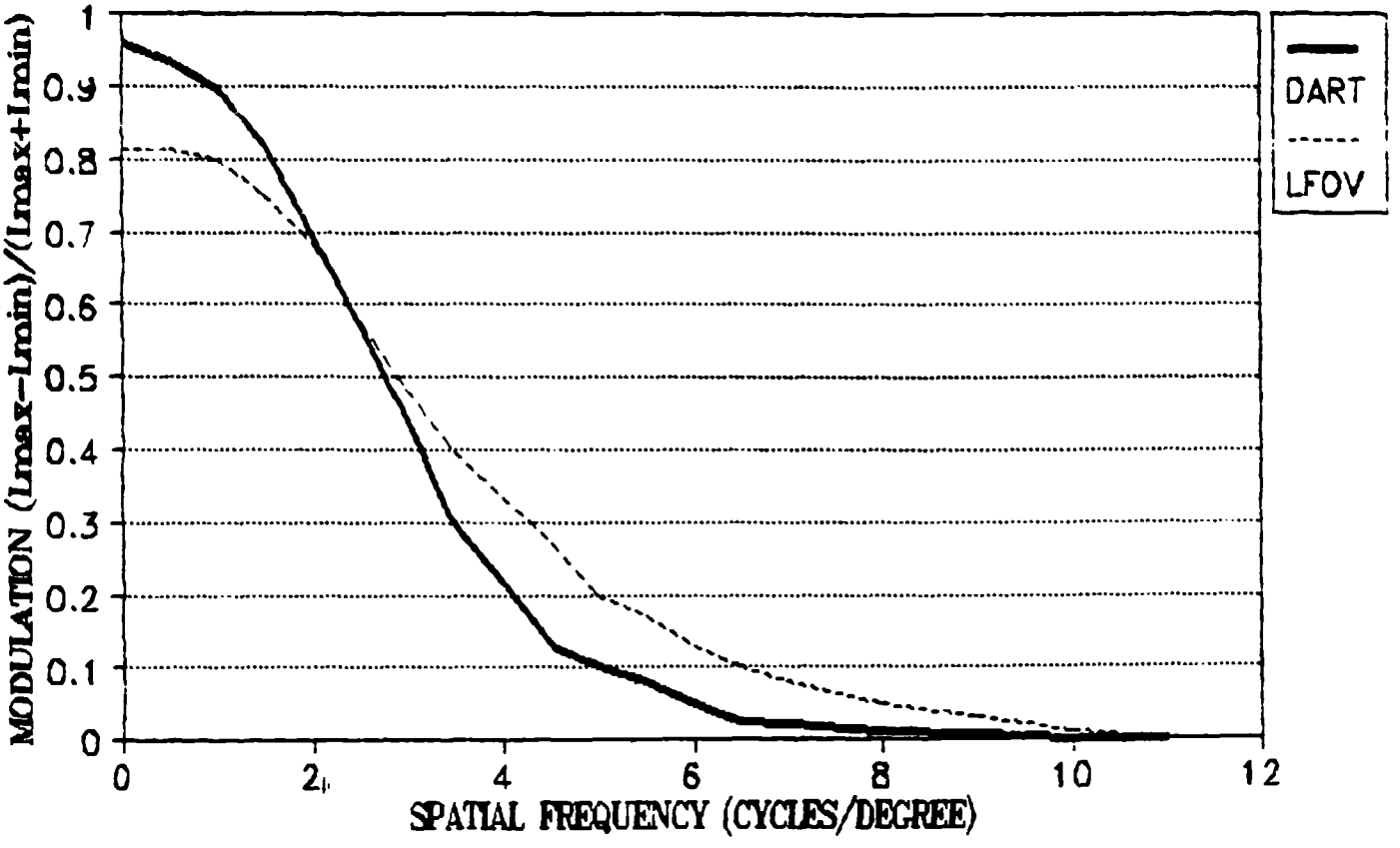


Figure 2
Michelson Contrast for DART and LFOV Displays

provides better contrast at spatial frequencies less than 3 cycles/degree of visual angle. The LFOV provides better contrast at spatial frequencies greater than 3 cycles/degree of visual angle, suggesting better resolution in the LFOV (assuming equal luminance capabilities).

Empirical results obtained by casual reports from observers suggest the DART is preferred to the LFOV. Most likely, though, this preference is a result of the luminance in the DART being approximately an order of magnitude greater than the LFOV. From a practical standpoint, it will be unlikely that the image quality of two display devices can be compared based upon changes in only one parameter (e.g., the display MTF). An image quality metric must then be able to integrate information across all these dimensions (e.g., low frequency contrast, resolution, luminance) in order to be valid across a wide range of display devices. Answers to image quality questions, such as that posed in Figure 2, will require a multidimensional approach to image quality research.

Visual displays employed in flight simulators represent a subset in the display domain with unique characteristics pertaining to image requirements. For many training applications in flight simulators, it is crucial that the field of view be as large as possible. In many instances, this translates into maximizing the physical size of the display, which also indirectly affects other primary display parameters, including luminance and resolution.

The traditional approach to determining image quality, dating back to Schade (1956) and previously to the Strehl ratio (see Chapter 2 of Biberman, 1972), has focused on the MTF as the primary driver in development of an image quality metric. In this initial report, we will explore this premise (i.e., the use of only the display MTF as a predictor of image quality) more closely. The proposed methodology will consist of filtering static, achromatic images with mathematically derived display MTFs. Empirical observation and comparison of the resulting images may then be compared with metric predictions.

The logic above directs our attention only to the display MTF as a predictor of image quality. The display MTF may be used as an indicator of the quality of the display, independent of the original image and observer characteristics. Realistically, however, the display MTF characteristics are important only in the determination of image quality to the extent that the original image and the human observer capture or weight the same information. For example, a display device that can pass high spatial frequency information is of little practical use if the original images consist of information below 12 cycles/degree of visual angle. Toward this purpose, this report also identifies both image and observer visual system characteristics. These characteristics are analyzed with respect to their effect on the display MTF. In the next section, the rationale for use of the MTF

in image quality prediction is developed through the introduction of a Linear Systems approach.

DISPLAY MODULATION TRANSFER FUNCTION AND LUMINANCE MODULATION: A LINEAR SYSTEMS APPROACH

Linear Systems Matrix Representation

The luminance profile of a visual display along one spatial dimension (vertical or horizontal) can be represented as a waveform which varies continuously in luminance as a function of spatial location (e.g., Y(x) where x denotes location and Y denotes the luminance at that location). This waveform can be mathematically represented in the frequency domain by the Fourier Transform of Y(x):

$$F(f) = \int_{-\infty}^{+\infty} Y(x)\, e^{(-2\pi i f x)}\ dx. \qquad (1)$$

Equation 1 is simply a decomposition of the luminance profile, Y(x), onto a set of basis functions, denoted by $e^{(-2\pi i f x)}$. In this instance, the basis functions redistribute the one-dimensional waveform in the spatial frequency domain. A simple analogy is that our position in space can be represented in three dimensions (three basis functions). These basis functions are not unique and many alternative ways of representing the positional information exist.

In a discrete or digitized version of the luminance profile representation, such as would be used to represent digital-to-analog codes (DAC) on a raster line of a display, the luminance values could be represented by a vector $\langle Y_{nT} \rangle$ of length N. Here n denotes the nth position on the raster line with a spacing of T (in the desired units of visual angle or linear distance between elements on the screen). The corresponding frequency representation would be:

$$F(k\omega) = \sum_{n=0}^{N-1} Y_{nT} e^{-j\omega Tnk} \quad for \quad k=0,1,\ldots N-1 \qquad (2)$$

where $\Omega = 2\pi/NT = 2\pi f$ is a measure of frequency separation between the samples. We denote the discrete Fourier Transform of the vector $\langle Y_{nT} \rangle$ n=0,N-1 by the vector $\langle F_{k\Omega} \rangle$ k=0,N-1 and k represents a position on the spatial frequency axis. If the raster line subtended X degrees of visual angle, the center-to-center distance between individual elements or pixels would be X/N in degrees of visual angle where N represents the number of elements in the

vector. In the frequency domain, the theoretical highest spatial frequency represented would be N/(2X) cycles per degree of visual angle because a cycle requires an on-off sequence. The frequency separation between individual spatial frequency components would be 1/X cycles per degree of visual angle and there would be N/2 positive spatial frequency components and N/2 negative spatial frequency components.

To approximate the luminance modulation across the width of the field, the frequency-based representation may be used with a number of the high frequency components left out of the signal (i.e., a low pass filtering process). This procedure is equivalent to performing polynomial regression (see Fig. 3) where complex functions are approximated by polynomials which typically contain fewer parameters, and thus can be represented more efficiently. This data compression technique is one advantage of using Fourier analysis to represent signals in a dual space mode.

Linear systems analysis provides a computationally efficient method of representing and predicting signals which have undergone linear (or approximately linear) transformations. A linear transformation occurs when the output, Y_i, is simply a linear combination of the input,

$$Y_i = \sum_{j=1}^{N} a_j x_j \qquad (3)$$

where x_j represents the input and a_j represents the coefficients in the linear combination. The input in our case is the two-dimensional static luminance profile. In the discrete case, then, we assume that the output luminance for a single pixel is a linear combination of the input (e.g., DAC values) at individual pixel locations.

In a truly linear system, the coefficients are independent of other parameters (e.g., location on the visual display, display luminance), but true linearity is rare in practice. For example, no practical display devices are truly homogeneous in luminance, contrast, resolution, and color across the entire display. In many instances, though, approximate linearity is assumed and system coefficients (i.e., a_j) are approximated by using measurements from the center of the display or by averaging across measurements taken at different locations from the display.

As an example of a linear system, let us suppose that we wish to represent the luminance output of a pixel (Y_{ij}) of a display from the DAC values of the pixels (x_{ij}). If we assume that the transformation is approximately linear (e.g., the voltage gamma of

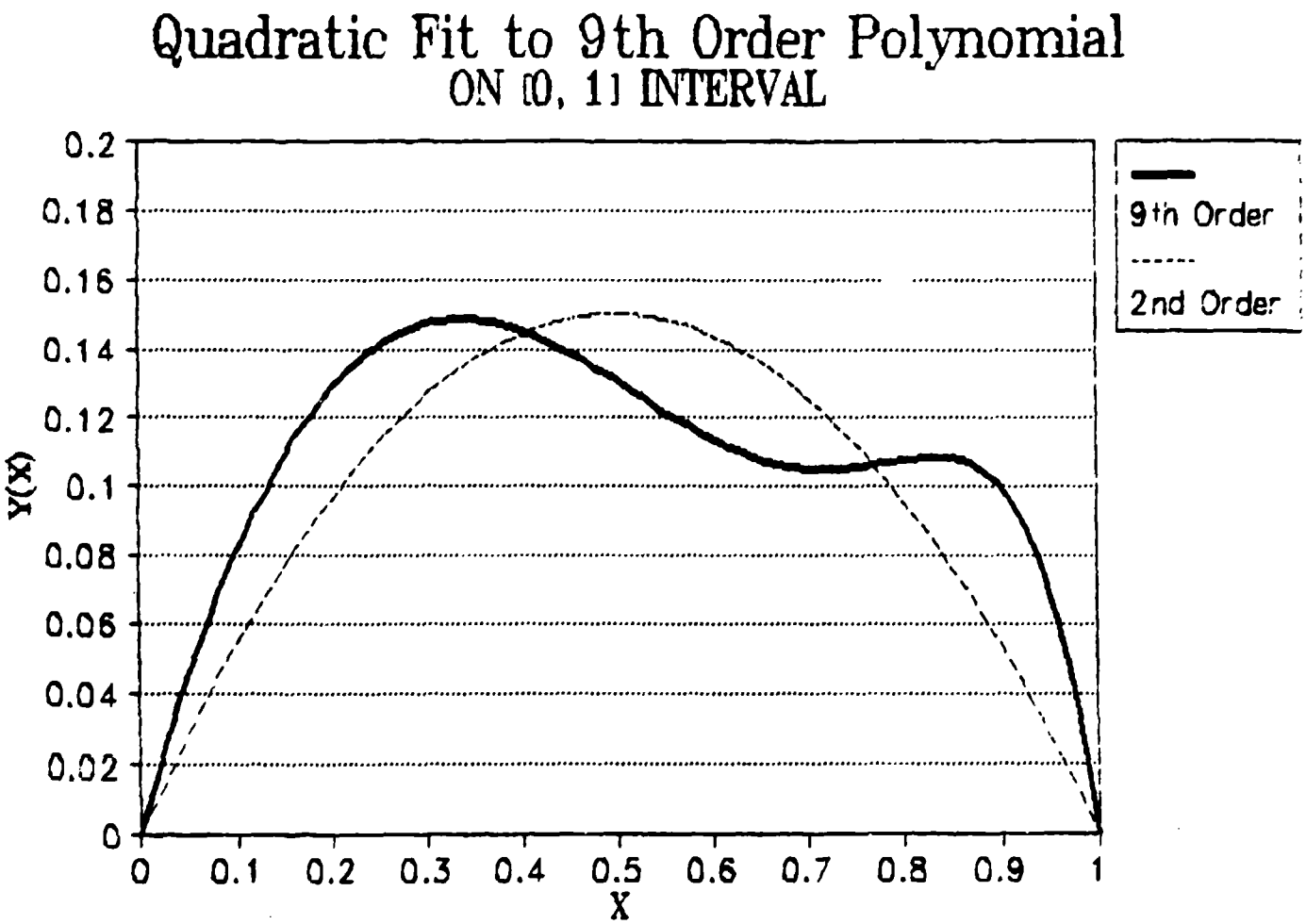


Figure 3
Low Pass Spatial Frequency Approximation

the display, spatial homogeneity) and that only the adjacent neighbors of x_{ij} have any effect on the luminance output of Y_{ij}, then the linear relationship between luminance output and DAC input is as follows:

$$Y_{ij} = \sum_{k=i-1}^{i+1} \sum_{l=j-1}^{j+1} a_{kl} x_{kl} \tag{4}$$

Based upon this equation, only nine nonzero coefficients are required and these can be represented by the matrix A such that:

$$A = \begin{bmatrix} a_{-1,-1} & a_{-1,0} & a_{-1,1} \\ a_{0,-1} & a_{0,0} & a_{0,1} \\ a_{1,-1} & a_{1,0} & a_{1,1} \end{bmatrix} \tag{5}$$

and y_{ij} and x_{ij} can be represented by the matrices

$$Y = \begin{bmatrix} y_{ij} \end{bmatrix} \quad X = \begin{bmatrix} x_{ij} \end{bmatrix}. \tag{6}$$

We now say that Y is the convolution of A with X or that Y = A*X. In the frequency domain, however, the convolution operation becomes a simple multiplication such that if F(A), F(X), and F(Y) denote the discrete, complex valued, Fourier transform of A, X, and Y, then F(Y) = F(A)F(X). Y may then be recaptured through the inverse Fourier transform $Y = F^{-1}[F(Y)]$. The energy in the input and output signals is denoted by the squares of the amplitudes of the signals. For complex valued functions, these squares are obtained by multiplying the signal (F(Y) or F(X)) by their complex conjugates ($F^*(Y)$ or $F^*(X)$). The energy in the output signal, Y, is denoted by:

$$[F(Y)F^*(Y)]^{\frac{1}{2}} = |F(Y)| = |F(A)||F(X)|. \tag{7}$$

The Modulation Transfer Function of this one-dimensional signal is defined as:

$$MTF = |F(A)| = \frac{|F(Y)|}{|F(X)|}. \qquad (8)$$

An additional assumption in this linear systems approach is that energy is conserved between input and output. The result of this is that:

$$|F(Y)|_{f=0} = |F(X)|_{f=0} \quad or \quad |F(A)|_{f=0} = 1 \qquad (9)$$

or the MTF is identically 1 at zero spatial frequency (i.e., f = 0) or DC (Direct Current) frequency. If the system contains some gain or loss (e.g., < 100% transmission of a lens), the modulation amplitude at zero frequency will not be identically 1. As will be discussed in the following section, luminance modulation on a display device at zero or DC spatial frequency will never be identically 1. Interpretations of display MTFs which normalize the modulation curves will be shown to be ambiguous.

The previous explanation is a brief introduction to linear systems and the MTF of a linear system. This linear systems approach is employed in representing luminance contrast. With display devices, the device's capability of maintaining contrast with increasing fineness of detail is a major contributor to image quality. Defining and measuring contrast, however, is not a straightforward process. Various measures of contrast, their relationship to one another, and their relationship to other display properties are discussed next.

Luminance Modulation: Contrast Measures

Contrast is a measure of relative luminance variability defined for some spatial extent. In the visual system, neural mechanisms such as excitatory and inhibitory center-surround nets are physiological evidence of contrast-related functions in the visual hierarchy. In statistical entities such as a Normal Probability Distribution, moments such as the mean and variance are independent of one another. In visual perception, however, sensitivity to contrast does change as a function of mean luminance.

Although more definitions of contrast exist, three operational definitions of contrast are commonly used in the literature. They are as follows:

$$(a)\ C = \frac{L_t}{L_b} \qquad (b)\ C_R = \frac{L_t - L_b}{L_b} \qquad (c)\ C_M = \frac{L_t - L_b}{L_t + L_b}. \qquad (10)$$

L_T and L_B represent maximum and minimum luminance, respectively. Equations 10a and 10b have been popular in display manufacturer guidelines and human factors work, respectively. Equation 10c has been popularized by the physiological literature and hypothetically relates a center-surround mechanism (e.g., Rogowitz, 1983) where discrimination of signals is given by the difference in output weighted by the average output (or twice the average) of the local area. Equation 10c is also typically called modulation depth as well as contrast. By measuring modulation depth of a sinusoidal waveform on a display for a range of spatial frequencies, a modulation depth curve is obtained. In most image quality work, a modulation depth curve which is normalized to a value of 1 at zero spatial frequency is referred to as the MTF of the display.

Note that, up to this point, no mention has been made as to the spatial extent of any of our contrast measures. For a complex image, then, the brightest luminance may occur at the edge of the image and the lowest luminance may occur in the center. With respect to the human visual system, defining the contrast between these two points has very little meaning in any objective or performance-related sense.

The Michelson Contrast or modulation depth was originally defined for a sinusoid waveform where, for the most part, the complete luminance cycle was contained in a local area. Peli (1990) provides an introduction to contrast in complex images and makes the argument that we should concern ourselves with localized measures of contrast.

For the three contrast measures presented here (i.e., C, C_R, and C_M), it is useful to note that the three equations are monotonically related to one another within a range of luminance values. For example,

$$C - 1 = \frac{L_t}{L_b} - 1 = \frac{L_t - L_b}{L_b} = C_R \quad and \quad \frac{C_R}{C_R + 2} = \frac{L_t - L_b}{L_t + L_b} = C_M. \quad (11)$$

C and C_R are affine transformations of one another (i.e., a linear transformation plus a constant) but are not linear transforms of one another from a strict linear systems viewpoint. The relationship between the Michelson Contrast, C_M, and the other two contrast measures is nonlinear. Any analysis of image quality using one measure of contrast cannot be assumed a priori to hold when using the alternative measures.

Given the monotonicity between the contrast measures within defined ranges of L_{max} and L_{min}, many ordinal comparison findings will translate across measures. In fact, for performance measures, a

range of useful contrasts may be derived as a function of the task. For example, Figure 4 is a rearrangement of target (a small circle) detection data obtained by Blackwell (1946). He experimentally determined size detection thresholds for circular targets as a function of contrast ($C_R = (L_T - L_B)/L_B$) and background luminance. From the data in Figure 4, we can estimate that (a) increases in background luminance yield little improvement in performance beyond approximately 11 fL of background luminance, (b) there is little improvement in detection performance for $C_R > .3$, and (c) performance is asymptotic for C_R around 1. Such data combined with equivalences provided in Equation 11 allow us to estimate that when working with Michelson Contrast, our range of interest is $0 < C_M < .86$ in accounting for performance variability in detection tasks. This analysis pertains well to detection tasks but is not necessarily relevant to suprathreshold tasks.

Michelson Contrast is the typical unit of measurement for empirical development of MTF or modulation depth curves. Modulation depth or MTF curves are plots of the Michelson Contrast as a function of the spatial frequency of a sinusoidal waveform displayed on a device. At zero spatial frequency, the Michelson Contrast is $(L_{MAX}-L_{MIN})/(L_{MAX}+L_{MIN})$ where L_{MAX} and L_{MIN} are the maximum and dark luminances of the display, respectively. Note that for spatial frequencies above one cycle per degree of visual angle, the Michelson Contrast will be an estimate of contrast localized within an area of one degree of visual angle. This is a spatial area which exists well within the fovea and we may reasonably assume this to be a localized contrast.

Many manufacturers provide a maximum-to-minimum (i.e., L_{max}/L_{min}) contrast ratio which can also be used to estimate the DC Michelson Contrast or the Michelson Contrast on the Y-axis. For example, with large projection displays, a 10-to-1, bright-to-dark contrast ratio is quite good. This would translate into a (10-1)/(10+1) = .82 Michelson Contrast Value. The .82 contrast or modulation depth, though, is only for low spatial frequencies. It provides no information concerning available contrast for target subtending only a few minutes of visual angle. Small helmet-mounted displays are able, in some instances, to generate 100-to-1 contrast ratios. This translates into a .98 Michelson Contrast at the DC or zero frequency level.

Because of the actual environment within which the display resides (e.g., a large dome), ambient lighting may contribute to manufacturer specifications for L_{MAX} and L_{MIN}. This can be approximated algebraically by adding a constant (the ambient illumination reflected off of the display) to the maximum and minimum luminance values. Ambient illumination will always lower the Michelson Contrast and requires us to consider the environment immediately surrounding the display.

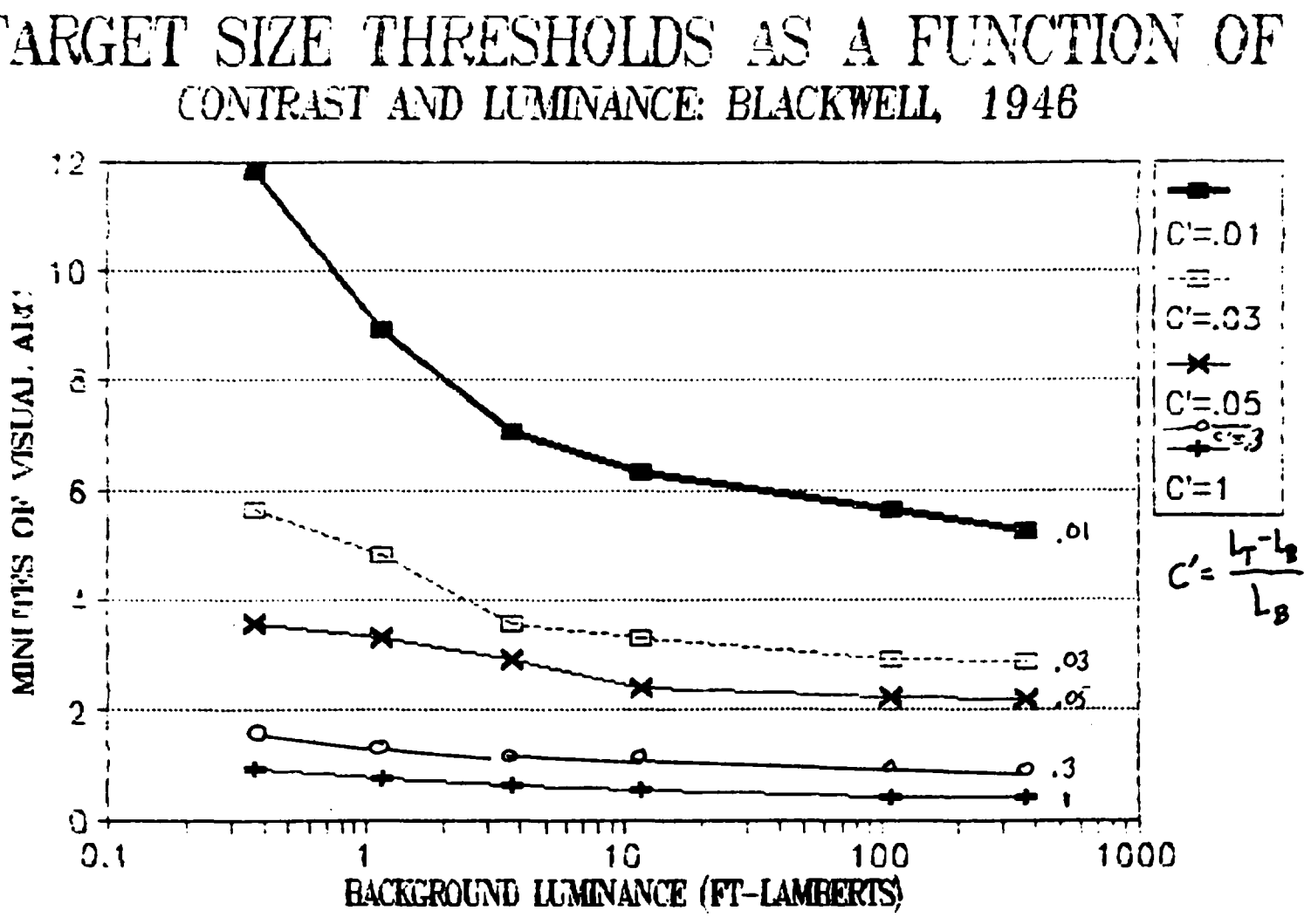


Figure 4
Detection Curves Estimated from Blackwell Data (1946)

One of the ambiguities discussed later in this report concerns the normalization of the modulation depth curves so contrast is identically 1 at zero spatial frequency. This normalization requires that the Michelson Contrast, or modulation depth, at all frequencies be divided by the modulation at zero or DC frequency. This normalization allows the contrast or modulation depth curve to appear as an MTF. For image quality purposes, normalization of these curves renders them ambiguous as parameters in a metric for comparison with other displays simply because the normalization coefficients differ across displays (Evans, 1990).

The two methods (direct and indirect) of empirically generating modulation depth curves or MTFs are discussed in the next section. The indirect method generates an actual MTF curve and, if this curve is used to represent Michelson Contrast or modulation depth, it must be unnormalized (i.e., multiplied by the modulation depth for zero frequency at all spatial frequencies). The following section is not meant to be a technical overview of MTF measurement but is included to show problems inherent in the methods of measurement which can have serious consequences on their use in metrics. For a more comprehensive review of MTF measurements, see Kelly (1992), Veron (1985), or Beaton (1989).

Measuring Luminance Modulation From Display Devices: Direct Versus Indirect Methods of Measurement

The physical size of displays used in flight simulation may vary from a helmet-mounted display (HMD) on the order of a few square inches to a full field-of-view, 24-foot diameter dome which covers a surface area of approximately 600 square feet. As expected, the ability to project a luminance profile onto these media can easily vary by a factor of 100 (e.g., 1/2 fL on some large domes to 50 fL on an HMD). Along with the luminance capabilities, the observer viewing distance also varies widely across display device. For perceptual purposes, the observer viewing distance is critical to any analyses as this helps determine the actual size of the retinal image.

For a display which extends a linear distance, L, in the vertical or horizontal direction, and a distance, d, from the observer, the visual angle subtended in the respective direction is:

$$\tan\left(\frac{\theta}{2}\right) = \frac{L}{2D} \quad or \quad \theta = 2\tan^{-1}\left[\frac{L}{2D}\right]. \qquad (12)$$

Mathematically, when the argument of the tangent function is quite small, the approximation tan(x) = x holds where x is measured in radians (2π radians = 360 degrees or 57.3 degrees/radian). In Equation 12, this simplifies to:

$$\tan\left(\frac{\theta}{2}\right) = \frac{\theta}{2} = \frac{L}{2D} \text{ radians or } \theta = \frac{L}{D} \text{ in radians.} \quad (13)$$

Converting Equation 13 into degrees of visual angle yields:

$$\theta = \frac{57.3L}{D} \text{ degrees of visual angle.} \quad (14)$$

Thus, for small visual angles, changes in the observer distance to the viewing screen are linearly related to changes in the visual angle subtended.

Equation 13 provides a standardized approximation for estimating the visual angle subtended by an image in the real world relative to the observer's eye. The visual angle subtended on the retina from the human lens (θ_{ret}) is not the same as the visual angle subtended on the retina by the actual image (θ_{im}) (Westheimer, 1986). A better approximation is $\theta_{ret} \approx .82\theta_{im}$. This fact does not change relative observations concerning real-world images and their analyses. However, it is pertinent when a discussion of neural processes (e.g., receptor spacing, visual cortex mapping) comes into play.

Contrast functions, modulation depth curves, or MTFs are usually reported or displayed (i.e., the unit on the x-axis of a graph) as a function of spatial frequency in cycles per degree of visual angle subtended from the user's viewpoint. Less frequently, these functions may be reported as a function of linear distance on the actual viewing device or distance on the retina in millimeters. Reporting them as a function of visual angle allows direct comparison to visual system functions from the human (e.g., the CSF). The drawback, though, is that the observer-viewing distance is required for calculation, as shown in the previous paragraph.

If image quality is to be determined from the observer's viewpoint, the observer-viewing distance should always be considered an integral part of the viewing apparatus. It is plausible to imagine two display systems such that at a viewing distance, D_1, System A provides better image quality than System B, but at distance $D_2 > D_1$, there is no difference between the two systems. Hypothetically, System A might provide noticeably better high spatial frequency contrast than System B. As the viewer distance from the displays increases, however, the contrast curves are shifted toward the higher end of the spatial frequency curve and the contrast improvement of System A over System B may be beyond the resolution limit of the human visual system. In the

limit, this argument is true for all display systems. As the viewing distance approaches infinity, there will be no image quality difference between displays.

In order to measure the contrast or modulation depth from a display, two methods are typically used, the direct and the indirect methods (see Beaton, 1989, or Kelly, 1992). In the direct method, a luminance-varying sinusoidal waveform of a specified spatial frequency is displayed on the viewing device. The peaks of the input waveform should be set for the minimum and maximum capabilities of the display device. A photometer is used to measure the peaks, L_{max}, and valleys, L_{min}, of the luminance profile directly off the viewing device for a discrete number of spatial frequencies. For each data point, $C_M = (L_{max} - L_{min})/(L_{max} + L_{min})$ is plotted as a function of the spatial frequency, generating a curve. To obtain modulation at zero or DC frequency, the maximum luminance and dark field or minimum luminance are also measured from the display and plotted as C_M.

Because of the discrete nature of addressing many displays, it may be preferable to measure modulation using a squarewave input generated from pixels instead of sinusoids. The following approximation from Schade (1987, p. 6) may then be used to estimate the sinusoidal response from square wave responses:

$$MTF(f) = \frac{\pi}{4}\left(r(f) + \frac{1}{3}r(3f) - \frac{1}{5}r(5f) + \frac{1}{7}r(7f) + \frac{1}{11}r(11f) - \frac{1}{13}r(13f) \right). \qquad (15)$$

In Equation 15, MTF(f) is the MTF or sine wave response at f cycles per unit of distance (e.g., linear distance or visual angle subtended at the eye) and r(f) represents the square wave response at f cycles of the square wave per unit of distance. In addition to the eye's inability to respond to higher spatial frequencies, the response of the display also decays as the frequency increases. Thus, the higher powers in the square wave response in Equation 15 tend to zero and are dropped from the approximation. For any spatial frequency, f, the modulation resulting from the square wave will be an upper limit on MTF(f), the sine wave response.

The indirect method of measurement, as opposed to the direct method, requires only a single measurement (in time) to generate the modulation depth curve. In this method, a one-dimensional transform may be estimated by displaying a single line (e.g., illuminating a row of pixels) at maximum luminance on the viewing device. A spatial photometer is used to measure the luminance transition from dark to light across this element. The Fourier transform of this one-dimensional waveform represents modulation depth over the spatial frequency axis. Using this method,

modulation at zero frequency will be identically 1. All modulation values must subsequently be multiplied by the normalizing coefficient $(L_{max}-L_{min})/(L_{max}+L_{min})$ where L_{max} and L_{min} are the maximum and minimum luminance values available from the display device.

There are numerous technical problems in using either procedure (the direct or the indirect) to estimate modulation depth curves. Conditions such as the maximum display luminance used and the uniformity of the display represent two such hurdles. For example, the shape of the modulation depth curve may vary as a function of the L_{max} value used. Using a maximum luminance value which nearly overdrives the limit of the display system will yield a higher contrast at the zero frequency or DC level but can also yield less contrast at high frequencies. The answer to such problems from the viewpoint of designing metrics is unclear. In practice, display engineers tweak the display system to improve subjective image quality. There is no analytical process which compares to this empirical practice from a metric perspective.

Luminance modulation characteristics of devices such as cathode ray tubes (CRTs) have been thoroughly studied (e.g., Infante, 1985, 1986; Barten, 1984, 1985). In these devices, luminance is typically generated when an electron beam(s) fall(s) on a phosphorous surface which emits photons. In many instances, a mask with holes in it constrains where the electron beam(s) may fall. The dead spots between the holes in the shadow-mask determine the "black" areas of the display, and the distance between holes in the mask denote the pitch of the mask.

Murch and Virgin (1985) present a good introduction relating the resolution and addressability of such displays. For such raster displays, it is clear that when the screen or portions of it are completely lit, viewers do not wish to see the individual raster lines. Therefore, the luminance modulation of the display at the frequency of the raster spacing must be below the observer threshold. On the other hand, it would be desirable that when alternate raster lines are lit, there is as much luminance modulation as possible. The two competing demands require a trade-off in display design.

The modulation depth curves (or MTFs) estimated from the direct and indirect measurement methods form the cornerstone of image quality metrics. Theoretically, then, it would be of interest to generate different display MTF shapes and test how these MTFs affect subjective or empirically measured image quality as well as the numerical metrics. In the next section, a mathematical formulation for estimating MTFs which allows us to move easily between two-dimensional space and the spatial frequency domain in two dimensions is introduced. In later sections of this report, this formulation will be used to generate a range of testable image quality scenarios.

Computation of Mathematically Derived MTFs

For work with digitized images in the study of image quality, it is useful to design computationally efficient filters to approximate the empirically obtained MTFs. Traditional Fourier transforms may be used toward this purpose and recent work in display measurement (e.g., Barten, 1984, 1985, 1988a; Infante, 1985, 1986) has shown the approximate relationship between display parameters (e.g., electron spot size) and computational formulas. Equation 16 below shows the correspondence between a point spread function in the spatial domain on the left and its Fourier transform on the right.

$$f(x) = a\sqrt{\frac{\pi}{b}}e^{-\frac{\pi^2x^2}{b}} \quad <=> \quad F(f) = ae^{bf^2} \tag{16}$$

The parameter "π^2/b" in the exponent of the point spread function (i.e., the leftmost equality in Equation 16) can be related to the distribution of current density in a CRT for a Gaussian spot (e.g., Barten, 1985). By setting $\pi^2/b = 12/d^2$ (or $b = \pi^2d^2/12$), where d denotes the width of the electron spot of a CRT beam (typically in millimeters) at which the Gaussian profile is at 5% of its maximum, Equation 16 is an approximation for a CRT display MTF. Note that in Equation 16, the units of d must be the same as the units of x, and the units of d and x are both the inverse of the units of f. In the spatial domain as well as with f in the frequency domain, d must always cancel with the units of x.

MTFs generated using the right side of Equation 16 can be used to approximate a variety of empirically measured MTFs. The function in the leftmost equality of Equation 16 represents a convolution filter in the spatial domain which can be used to filter an image in the spatial domain. In a later section, curves generated by the right side of Equation 16 will be used to approximate hypothetical display MTFs using specific values for parameters a and b. The left side of Equation 16 will then be applied to images as a convolution filter, simulating the process of the images being viewed through the hypothetical display.

When viewing filtered images on a display (e.g., CRT), a double-pass process (i.e., two sequential filtering processes) is being applied to the image. Not only does the original filtering affect the quality of the image but the viewing of the filtered image through a second display is equivalent to filtering the image again (i.e., a double-pass process). If our intention is to evaluate the effect of a specific MTF on the quality of images,

this MTF must be multiplied by the inverse of the viewing display (vd) MTF (i.e., $MTF_i(f) \times MTF^{-1}_{vd}(f)$). The resulting MTF would then be used in the image filtering process and viewing of the filtered image through the display device would then complete the double-pass process (i.e., $MTF_i(f) \times MTF^{-1}_{vd}(f) \times MTF_{vd}(f) = MTF_i(f)$).

One of the problems in applying the mathematical filters to actual images involves the premultipliers on both sides of Equation 16, a on the right side and $a(\pi/b)^{.5}$ on the left side. If the left side of Equation 16 is used to filter an image in the spatial domain, the energy from the filter (i.e., the integral) must evaluate to identically 1 in order to retain the same overall energy or luminance. In statistical terms, the left side of Equation 16 must be a probability density function (PDF). If the PDF is normal or Gaussian, it is required that a = 1 in the left side of Equation 16 in order that the PDF integrate to a value of 1 over the limits of $\underline{x}$. When a = 1 in the rightmost equality of Equation 16, the function will then evaluate to 1 when f = 0, i.e.,

$$\left[1 \cdot e^{-bf^2}\right]_{f=0} = 1. \qquad (17)$$

Unless the DC or zero frequency value of the MTF is identically 1, the area under the convolution filter in the spatial domain will not be identically 1, and the filter will either add or subtract energy from the signal. In filtering imagery, this corresponds to changing the average luminance of the image. Thus, in order to simulate modification of the display MTF without varying the average luminance of the image, a single restriction (i.e., a = 1) on Equation 16 is required. With this restriction, Equation 16 provides an efficient method of approximating actual display MTFs. As previously discussed, the spatial filter from Equation 16 can be used to filter digitized images, simulating the effect of viewing the images through the appropriate display devices.

The characterization furnished in Equation 16 provides a simple computational mechanism for moving between the spatial domain and the frequency domain. Given the MTF of an actual display, the rightmost equality of Equation 16 may be fit to the MTF curve (using b as a parameter) and images may be filtered in the spatial domain using the leftmost equality of Equation 16. The filtered image will be an approximation of how the image would appear in the display of interest. The drawback with this approach is that the y-axis intercept of the simulated MTF must be 1 (or very close to 1 for the approximation). Looking back to Figure 2, the comparison of the DART display with the dome display, it is apparent that the approximation will not suffice for some displays. Thus, this approach only simulates the effect of display MTFs on

imagery for a subset of MTFs whose y-axis intercept in the spatial frequency domain is quite close to 1.

The method provided in this section allows us to simulate the effect of a variety of display MTFs on static, achromatic imagery. Examples of other methods used to vary display MTFs are the defocusing of the electron spot in a CRT (Barten, 1987), variation of luminance and picture size (van der Zee & Boesten, 1980), changing the gamma characteristic of a display (Roufs, 1989), the addition of ambient light (Barten, 1988a), and the addition of noise (Barten, 1991). From these and other examples, the difficulty in manipulating display parameters and obtaining performance measures in a controlled, experimental environment becomes more apparent. The inability to manipulate specific display parameters while controlling other parameters is one of the biggest drawbacks in experimental studies of image quality, especially multidimensional investigations. Next, we turn our attention from characterizing the importance of the display device to a discussion of the final subsystem within the image quality systems framework--observer characteristics.

LUMINANCE TRANSFER CHARACTERISTICS OF THE HUMAN VISUAL SYSTEM

The systems approach to image quality (see Fig. 1) incorporates the human eye/brain system as the third and final filtering component. A metric representing image quality preference as determined by humans should be filtered or weighted accordingly. Research (see Hood & Finkelstein, 1986, for an overview) indicates that the human visual system filters are a function of: (a) spatial frequency, (b) average local scene luminance, (c) retinal eccentricity, (d) size of the test image, (e) wavelength or color, and (f) motion or temporal properties.

The complexity of representing the human eye-brain filter as a function in a 6-dimensional space necessitates that we simplify the filter. From a linear systems approach, an MTF surfaces as a logical approach for filtering the incoming luminance-varying signal. Measuring the MTF of the eye-brain subsystem is quite complex and involves many assumptions. Of the six factors noted above, the traditional MTF (of the optics of the eye alone) is measured only as a function of spatial frequency with all other factors held constant. However, individual MTFs may be measured for changes in each of the six listed parameters but would result in an even more complex task. It would also negate the simplicity, and in many instances the assumptions (i.e., a lack of linearity), afforded by the linear systems approach.

Traditionally, in image quality the filtering or weighting component used for the eye-brain subsystem is the CSF. The CSF at frequency u, CSF(u), is the inverse of the amount of contrast (Michelson Contrast) required to detect a sinusoidally varying

luminance pattern of frequency u. The human CSF, as does the MTF, also varies as a function of average luminance, retinal eccentricity, wavelength, test pattern size, and temporal properties (e.g., Glenn, Glenn, & Bastian, 1985). Measurement of the CSF also requires psychophysical procedures (e.g., ascending staircase, method of adjustment) in order to estimate individual data points of contrast thresholds. Measuring the MTF of the eye requires even more complicated procedures.

Note that the emphasis in this section is to compare the CSF function with the MTF of the eye as filters for image quality metrics. This specific comparison takes root from the idea that the CSF is a threshold measurement and the MTF is a suprathreshold measurement. The distinction relates well to the typical dichotomy in empirical image quality assessment, performance tasks which are intuitively associated with threshold behavior, and subjective preference which is more closely associated with suprathreshold behavior.

In evaluating the effect of employing a CSF or an MTF of the eye as a filter on an image quality metric, the effect can be assessed independently of any empirical task. That is, either function (CSF or MTF) can be substituted into an equation and the results compared. This comparison is one of the goals of this section.

As a final caveat, note that in employing the term MTF in this section, the term "eye" has closely followed. Data included in this section on the MTF of the visual system are only estimates of the MTF of the optics in the eye. These estimates do not include any information concerning processing which initiates at the photoreceptors and continues through the visual pathways in the cortex. Further discussion concerning this incomplete representation is provided in the MTF section which follows.

The Contrast Sensitivity Function (CSF)

CSF is the psychophysical weighting function of the human eye/brain system traditionally used in image quality metrics. To measure the CSF, 1-dimensional sinusoidal waveforms varying in luminance are presented to observers. The luminance variation in the waveform usually occurs in the horizontal, vertical, or oblique directions (see Westheimer, 1970). For a sinusoid of a specified spatial frequency, the CSF at that frequency is the inverse of the amount of Michelson Contrast, $(L_{max}-L_{min})/(L_{max}+L_{min})$, necessary to discriminate the pattern from a homogeneous pattern. By measuring the contrast threshold and computing the inverse at a variety of spatial frequencies, the points may be plotted as a function of spatial frequency. Figure 5 denotes two empirical CSF curves obtained for two average luminance levels (.1 cd/m^2 and 10 cd/m^2) that are replotted from van Meeteren and Vos (1970). The two CSFs

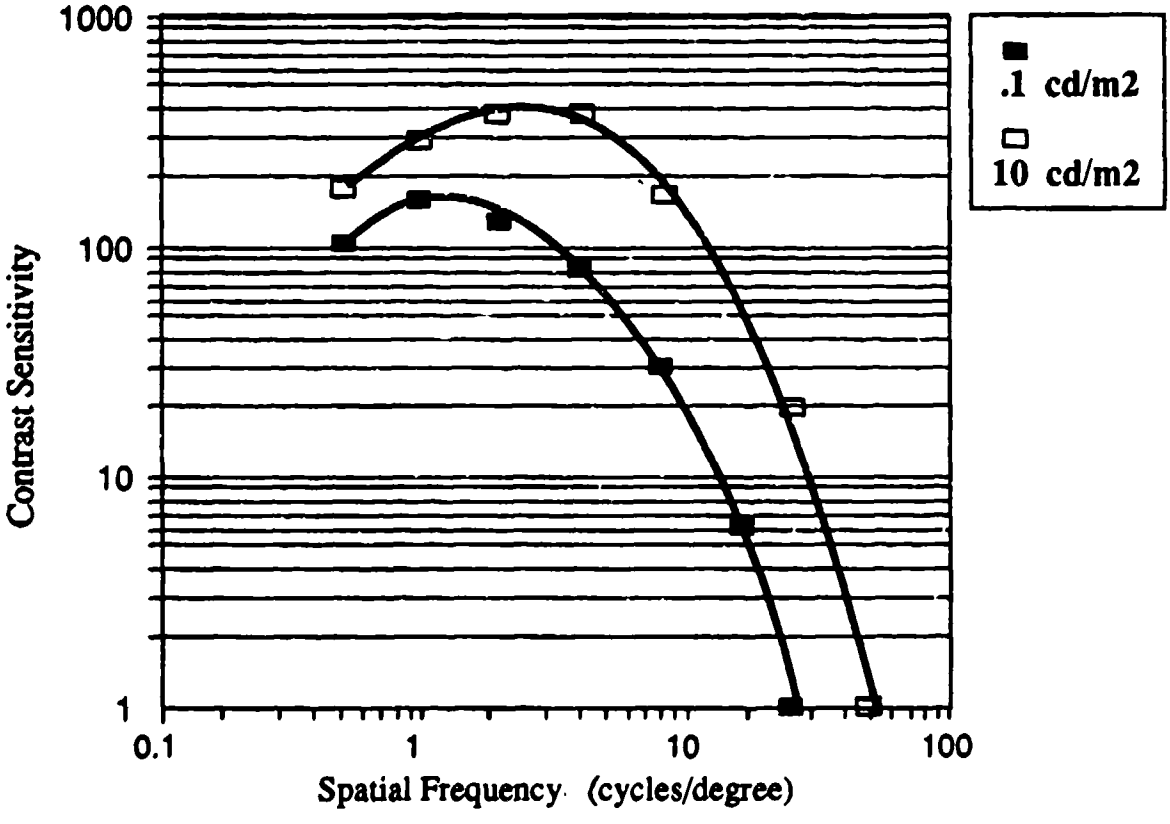


Figure 5
CSF Curves Replotted from van Meeteren and Vos (1972)

in Figure 5 represent data averaged over two observers using horizontal and vertical sinusoidal waveforms which covered a 17° X 11° field of view. Note that empirically measured CSFs are sensitive to many methodological parameters (e.g., test patch field of view, orientation of the pattern) as are most psychophysical thresholds.

With respect to image quality, the major point of interest in Figure 5 is the shape of the CSF. It appears as an inverted "U" function which typically reaches its maximum between 3 and 8 cycles per degree of visual angle, depending upon the average luminance of the waveform. If the CSF is used as a weighting function, midrange spatial frequency information will be emphasized relative to low or high frequencies.

It is interesting to note that the relative de-emphasis of low spatial frequencies is sometimes neglected. Infante (1985), for example, approximates the inverse of the CSF (measured at 10 cd/m^2) with the function $.0007655e^{.166f}$ (see Fig. 6) where f equals spatial frequency in cycles/degree of visual angle. This approximation neglects any loss in contrast sensitivity at low spatial frequencies.

The purpose of the examples shown here is to emphasize the use of the CSF or eye MTF in an image quality metric. In metrics such as the square root integral (SQRI), the CSF is a multiplicative weight. In models such as the Modulation Transfer Function Area (MTFA), the CSF is a subtractive threshold. In all applications, though, the capability of the display must be modified according to how the visual system will use the information. In addition, the manner in which the weight is applied should be conceptually interpretable.

The CSF has received more study than the MTF of the eye and this may be one reason why it gained popularity in use over the MTF of the eye. Many authors (e.g., Glenn, Glenn, & Bastian, 1985) have called the CSF the transfer function of the visual system. The CSF represents the psychological strength of the physical signal relative to the strength of the signal at other frequencies for the minimum detectable amount of contrast. The CSF may be the appropriate visual system filter to use in image quality metrics with detection tasks, but the MTF of the eye/brain system would be logically more appropriate for suprathreshold tasks. The fact that the shape of the transfer function would change from threshold to suprathreshold conditions violates the assumption of linearity as a function of scene luminance.

As mentioned, the human CSF is a highly variable psychophysical function. Not only does it vary as a function of

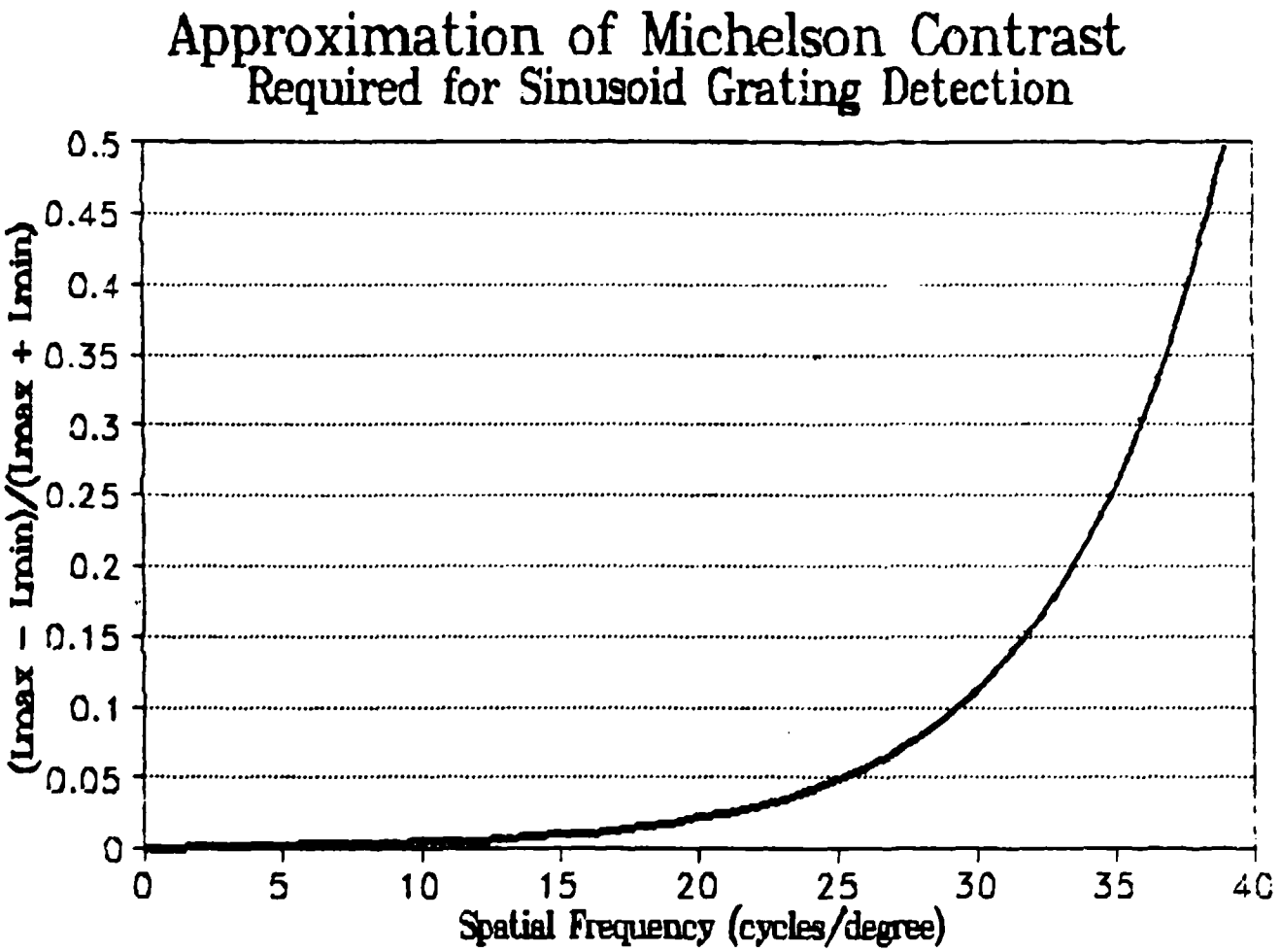


Figure 6
CSF Approximation Neglecting Loss in Sensitivity at Low Spatial Frequencies

stimulus duration, field size and orientation, temporal properties, retinal eccentricity, and wavelength, it is also highly variable across observers (e.g., sometimes by a factor of 10 at individual frequencies). For use in equations, then, the CSF is always an average of many observers and measured for a standardized setting.

For predictive use, there exists a number of mathematical approximations to the CSF. Beaton (1989) gives the following approximation:

$$CSF(F) = b_0^2 \exp[b_1 f + b_2 f + b_3 f] \quad (18)$$

where f is spatial frequency in cycles/degree of visual angle, b_0=.0017062, b_1=.2016188, b_2=-.0023161, and b_3=.0000002.

van Meeteren (1973) developed a numerical estimation for contrast sensitivity as a function of both spatial frequency and average luminance from data reported by van Meeteren and Vos (1972). Barten (1990) used empirical CSF data obtained by Carlson (1982) to include field size as a predictor of contrast sensitivity. The combined form of these efforts yielded the function:

$$CSF(f) = (A)(f)e^{-Bf}[1 + (c)e^{Bf}]^{.5} \quad (19)$$

where $A = 540(1+.7/L)^{.2}/(1 + 12/(w(1+f/3)^2))$, $B = .3(1+100/L)^{.15}$, C = .06, L denotes luminance in cd/m^2, w is the angular size of the display area calculated from the square root of the picture area, and f denotes spatial frequency in cycles/degree. Figure 7 shows a number of curves generated from Equation 19 with the field-of-view parameter, w, set to approximately 14°. For w > 10°, there is little variation in the CSFs. As the field of view grows smaller, sensitivity is lowered, or bows downward more, at the lower end of the spatial frequency spectrum.

In Figure 7, observe that peak sensitivity drifts from approximately 2 cycles/degree at mesopic levels of illumination to approximately 4-1/2 cycles/degree at photopic levels of illumination. This shift in peak sensitivity is a logical consequence of center-surround neural mechanisms at the ganglion layer of the retina. Within these mechanisms, spatial summation occurs in the center and the surrounding cells inhibit firing from the central, excitatory units. At lower luminance levels, lower frequency channels with excitatory components that integrate over larger areas of the retina are likely candidates for performing the detection process. Channels which match the light and dark regions of the sinusoid pattern may be implicated here. The excitatory portion of the center-surround mechanism could match the bright

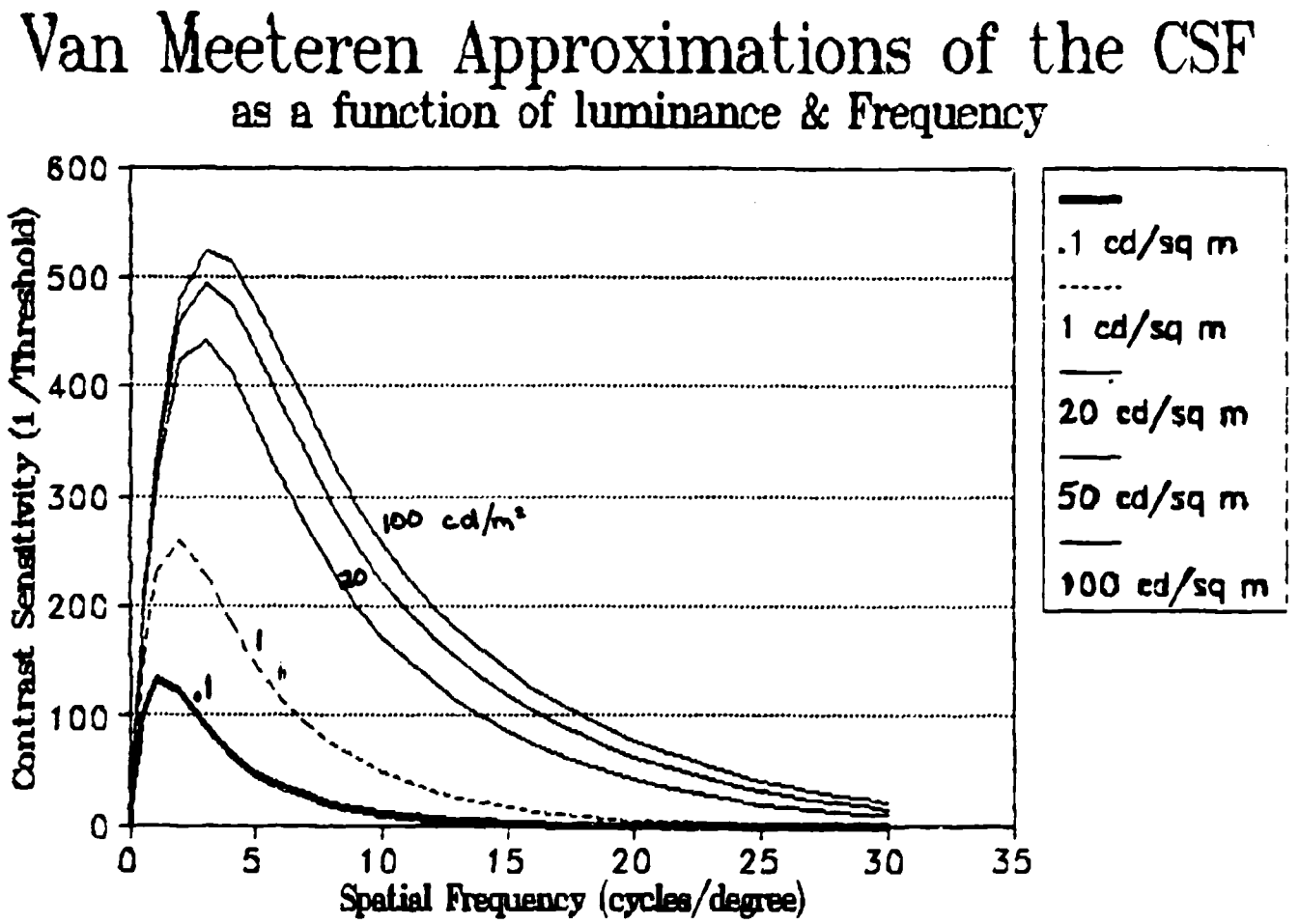


Figure 7
Van Meeteren Contrast Sensitivity Curves

portion of the signal and the inhibitory component of the center-surround mechanism falls upon the darker parts of the sinusoid, causing less inhibition in the channel output.

Many researchers (e.g., Campbell & Robson, 1968) suggest that the human CSF is an envelope of approximately seven independent spatial frequency channel analyzers (see Fig. 8). Carlson and Cohen (1980) and Carlson (1988) devised an image quality metric based upon just noticeable differences (JND) in Michelson Contrast for each of seven spatial frequency channels located at .5, 1.5, 3, 6, 12, 24, and 48 cycles/degree of visual angle.

Figure 9 shows an example for two representative display MTFs. The JNDs in Figure 9 are empirically obtained using sinusoidal waveforms and measurements of difference thresholds (for Michelson Contrast) at each of the seven spatial frequencies. In their model, the metric is the cumulative number of JNDs in all of the spatial frequency channels which lie below the display MTF.

In other metrics of image quality, the CSF is employed as a multiplicative weight in the spatial frequency domain. More contrast sensitivity at a given spatial frequency is used to imply that the human emphasizes this band of spatial frequencies more heavily in determining image quality. In such instances, when making ordinal comparisons of image quality across display devices, the CSF need only be unique up to a multiplicative constant (i.e., $CSF_1(u) = A \times CSF(u)$). Thus, if only an ordinal metric of image quality is desired, the CSF may be scaled up or down (i.e., normalized) to any constant.

Mathematically, this conjecture is as follows. Let two displays be represented by MTF_1 and MTF_2 with their respective metrics given as follows:

$$IQ_1 = \int_u G[CSF(u)\,MTF(u)]\,du \tag{20}$$

and

$$IQ_2 = \int_u G[CSF(u)\,MTF(u)]\,du. \tag{21}$$

In Equations 20 and 21, G is a monotonically increasing function of the product of the CSF and MTF where both the CSF(u) and MTF(u) are greater than zero for all u, and u is spatial frequency in cycles/degree of visual angle. If we multiply or scale the CSF functions

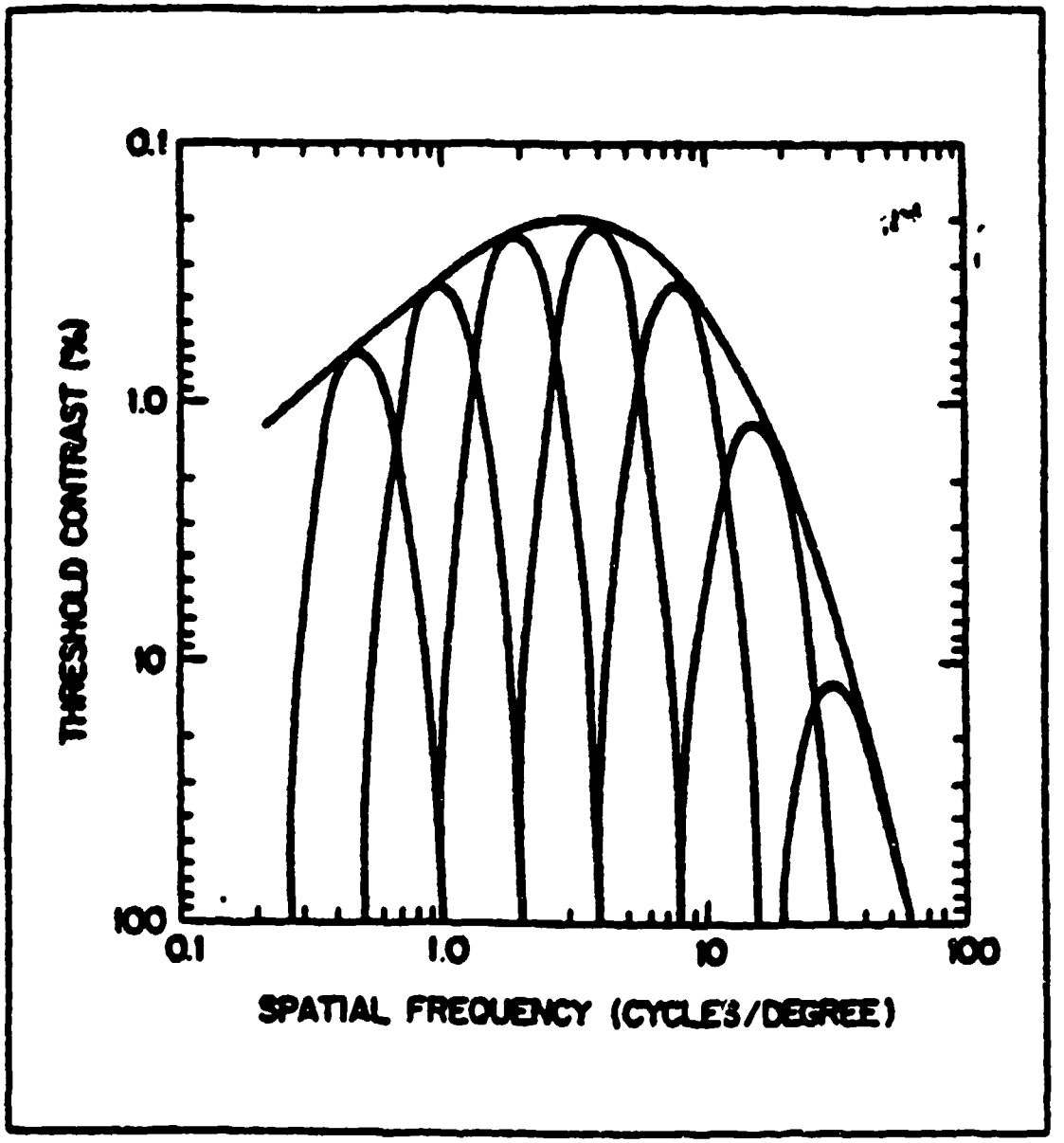


Figure 8
CSF as the Envelope of Seven Filters

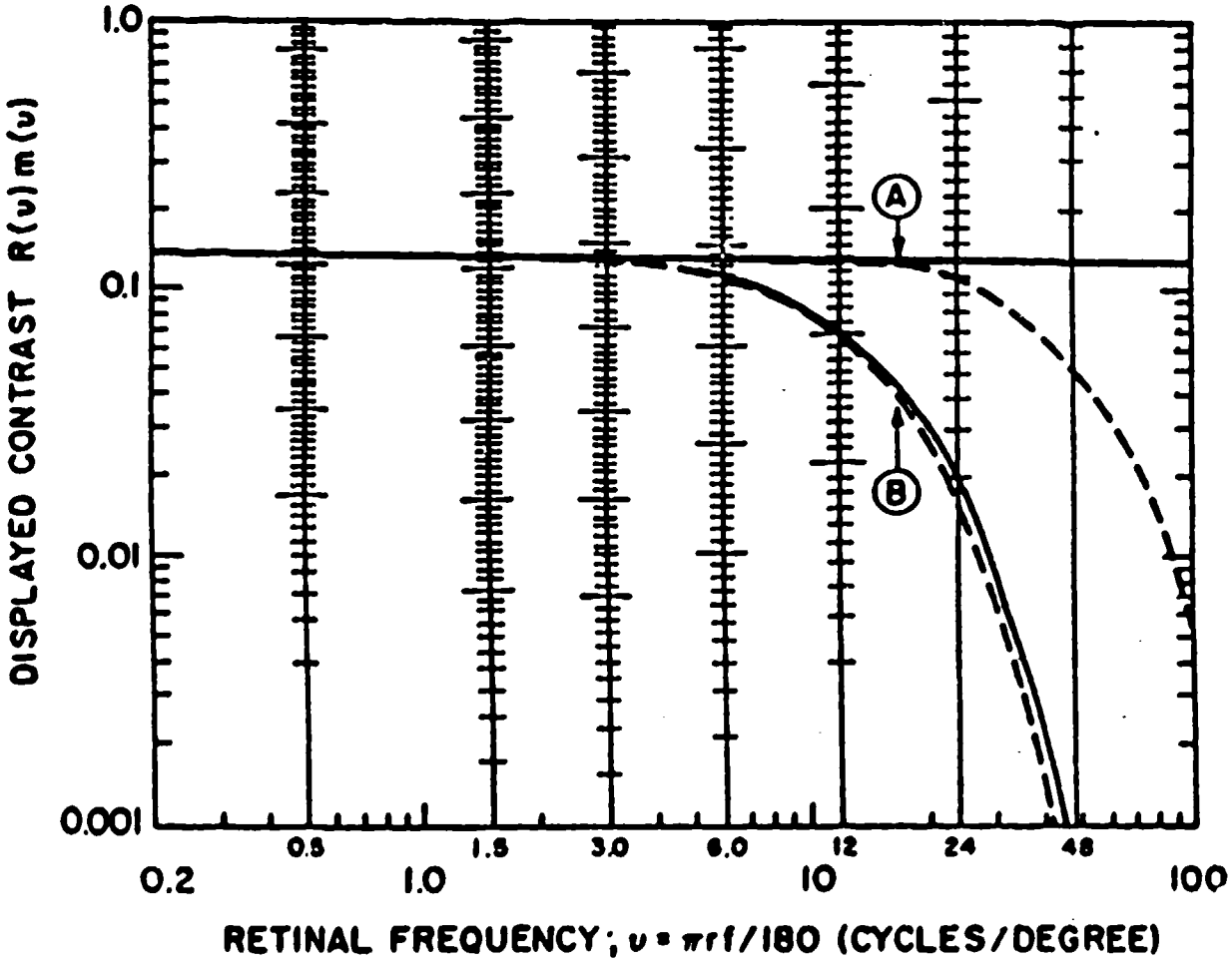


Figure 9
Carlson and Cohen JND-metric for a Display MTF
(Reprinted with permission,
courtesy Society for Information Display)

by a constant, C, at all spatial frequencies, the resulting image quality metrics are:

$$IQ_1^* = \int_{u=0}^{u_{max}} G[C \; X \; CSF(u)MTF(u)] \; du \qquad (22)$$

and

$$IQ_2^* = \int_{u=0}^{u_{max}} G[C \; X \; CSF(u)MTF(u)] \; du. \qquad (23)$$

For any ordinal ordering of IQ_1 and IQ_2, the same ordering also applies to IQ_1^* and IQ_2^*. For example, if $IQ_1 < IQ_2$, then $IQ_1^* < IQ_2^*$. This monotonicity across the scaling of the weighting factor allows us to scale the CSF for comparison with other multiplicative functions (e.g., the MTF of the eye). We can then look at their relative effect on image quality metrics.

The CSF approximation in Equation 19 permits us to weight image quality metrics from the observer viewpoint as a function of spatial frequency and average luminance. Other visual field parameters (e.g., retinal eccentricity, orientation) are held constant for the CSF measurement but will vary considerably and unmanageably in reference to general image quality. The CSF represents visual system sensitivity at a spatial frequency as the inverse of the contrast threshold at that spatial frequency. It is the relative transfer of physical quantities (contrast) across spatial frequencies required to obtain a specific psychological response (detection or discrimination). As many critics point out, though, the transfer function at the threshold level need not be equivalent to the transfer at suprathreshold levels. As should be pointed out, though, if the transfer function does change as a function of some variable (e.g., average field luminance), linearity is violated. The next section introduces empirical research used in estimating the MTF of the eye.

The Modulation Transfer Function (MTF) of the Eye

The linear systems approach posits that the complete system response of a linear system to an input in the frequency domain is the product of the individual transfer functions of the components or subsystems. This approach has popularized the notion of the MTF in general. In image quality work, critics have questioned why the CSF has been employed as a filtering function rather than the MTF of the eye/brain system.

This section develops the use of the MTF of the eye as an alternative to the use of the CSF as a weighting function in image

quality metrics. Of critical importance here is that the measurements discussed in this section include only the effects from the optics of the eye and nothing beyond this (e.g., neural mechanisms, transduction, cortical processing) in the visual pathway.

There is research suggesting that the neural system compensates for the MTF of the eye. Snyder and Srinivasan (1979) argue that for suprathreshold viewing conditions, neural processing compensates for optical degradation as seen in the MTF of the eye. For suprathreshold viewing conditions, they suggest a flat transfer function, a concept which when employed in a metric would mean essentially no compensation of the incoming signal by the human visual system. The basis for their argument comes from experiments which matched sinusoidal gratings of different spatial frequencies for their apparent contrast (Blakemore, Muncey, & Ridley, 1973; Georgeson & Sullivan, 1975; Kulikowski, 1975; Watanabe, Mori, Nagata, & Hiwatashi, 1968). At suprathreshold conditions, apparent contrast does not vary as a function of the spatial frequency of the sinusoid.

Implications from the argument above are that neither the CSF nor MTF of the eye are applicable filters, at least for suprathreshold activities in image quality. Empirically, there is little evidence suggesting that either filter represents what occurs from the perspective of the visual system and image quality. A plausible approach is to design and conduct empirical studies which are able to make predictions and test metrics based upon interchanging the filters. From this perspective, then, it is still reasonable to study how the use of the MTF of the eye in image quality metrics affects predictions about image quality.

Obtaining empirical estimates of the MTF of the eye is more complex than obtaining estimates of the CSF. First, note that the CSF is a psychophysical function for the complete visual system. The MTF, as presented here, is an estimate of the transfer function of the optics in the eye alone. In the past, researchers employing a linear systems framework were aware that processes in the visual pathway such as transduction were inherently nonlinear. Because the optics of the eye is an optical system (i.e., a system of apertures and lenses), a linear systems approach for describing the filtering of the eye was a natural approach. The following data represent only a development of the MTF of the eye and not of the entire visual system.

Campbell and Green (1965) and Campbell and Gubisch (1966) used interference fringes to form luminance-varying sinusoidal patterns of high contrast directly on the retina. Through measurements of the return or retinal reflection on this signal, they were able to estimate the MTF due to the optics of the eye. Figure 10 shows a family of MTF curves as a function of the spatial

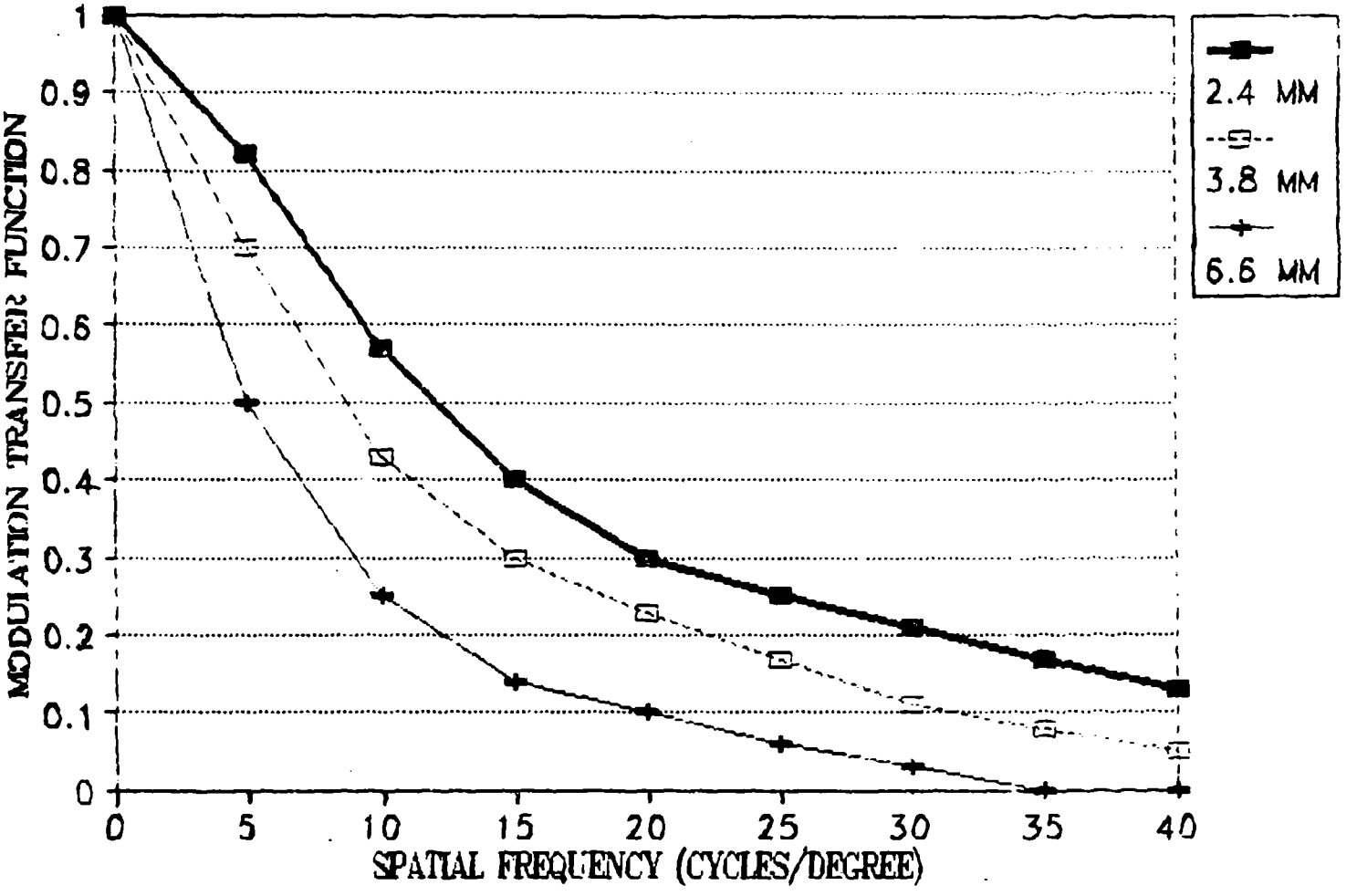


Figure 10
Gubisch (1967) Estimate of the MTF of the Eye

frequency of the sinusoid pattern and the pupil or aperture size. Note that as with any optical system, an increase in pupil or aperture size has the effect of decreasing resolvability.

Veron (1985) reported an estimate for the MTF of the eye which is as follows:

$$MTF(f_r) = e^{-.166 f_r} \qquad (24)$$

In Equation 24, f_r is the retinal spatial frequency and may be calculated as $f_r = (\pi Lu)/180$ where u is the number of cycles per unit of distance on the display and L is the observer distance to the screen measured in the same units as u. L and u may be bypassed, and $MTF(f_r)$ can be plotted as a function of f_r in cycles per degree of visual angle. Figure 11 shows such a plot of Equation 23 along with the MTF for a 6.6 mm diameter pupil from Gubisch (1967). The height and shape of two curves in Figure 11 closely resemble one another.

The Use of the CSF Versus the Eye MTF in Image Quality Metrics

Some researchers currently employ image quality metrics as if the numbers derived from them existed at the interval level. For example, consider Figure 12 as an ordering of four display devices (A, B, C, and D) on an interval-level metric. In Figure 12, assume the output of a metric is said to be in JNDs and devices A, B, C, and D receive scores of 10, 15, 20, and 27 JNDs, respectively. If the scales satisfy an interval level of measurement, device D is more preferable to device C than device B is to device A.

A more rational proposition at this time is to assume that image quality metrics satisfy only an ordinal level of measurement. In Figure 12, then, the only information available is that, in terms of preferences for devices, D > C > B > A where the "greater than" sign can be thought of as "is preferred to." Only after obtaining more complete knowledge of the weighting schemes used by the visual system should an attempt be made to extend the scale beyond an ordinal level of measurement.

If metrics are used as ordinal scales and the filter applied from the human subcomponent is a multiplicative weight, the filter may be scaled as suggested in the section on "The Contrast Sensitivity Function." Using these assumptions, it is not the absolute height of the weighting function or filter that matters but only the relative height across the spatial frequency axis. This allows us to scale both the CSF and the MTF of the eye to a height of 1 at their relative maximum and compare them on a single

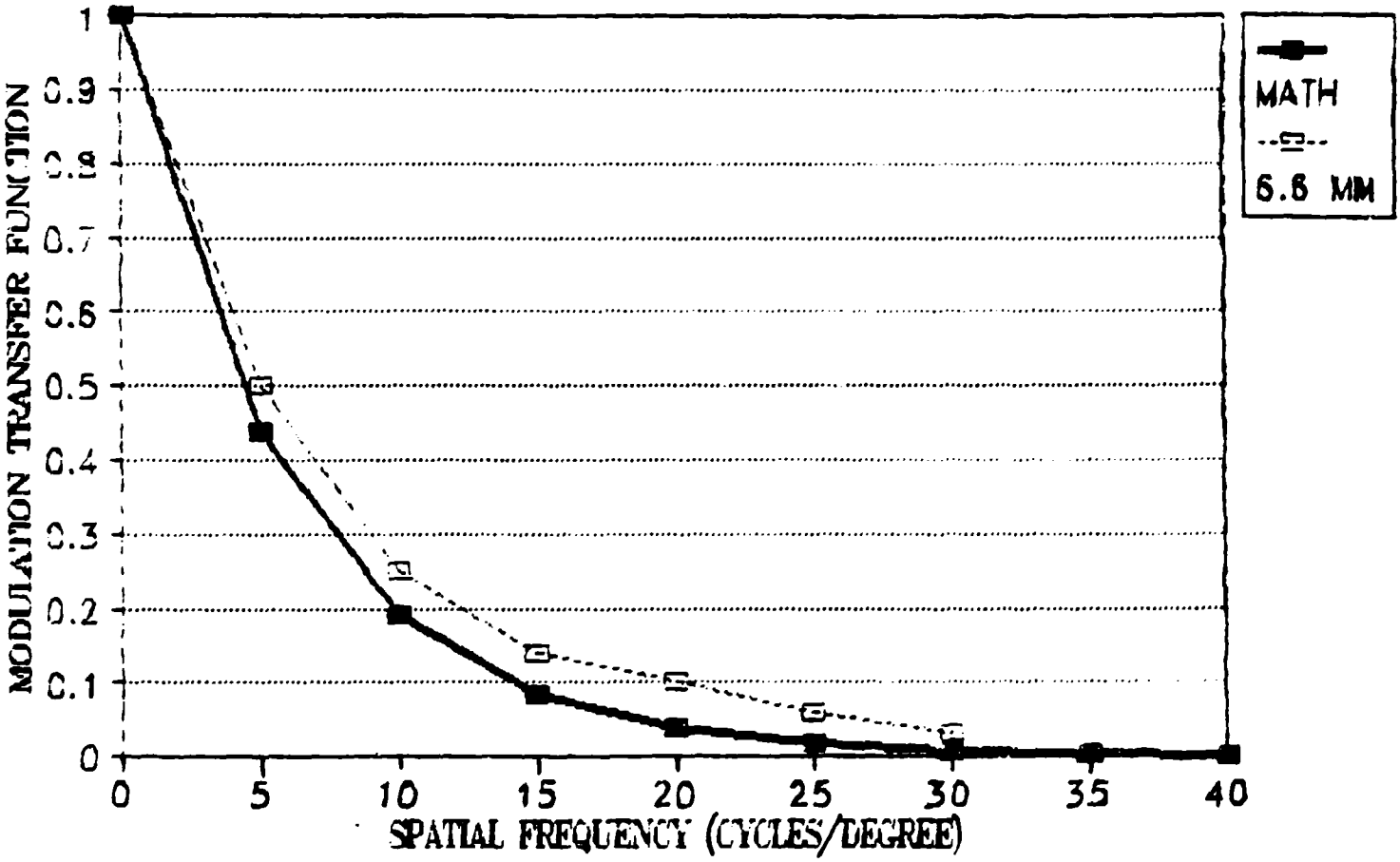


Figure 11
Mathematical Estimate of the MTF of the Eye

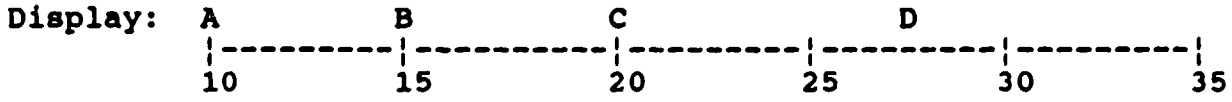


Figure 12
Metric Ordering on Interval-Level Scale
for Four Display Devices

graph. Figure 13 shows a comparison of Equation 24 (approximation of the MTF of the eye) with Equation 19 (CSF approximation) as a function of spatial frequency when the heights of the curves are normalized to 1. Figure 14 shows the same comparison as a function of the natural log of spatial frequency. Comparison of Figure 13 with Figure 14 shows the relative weighting of display information by spatial frequency when we change from a normal integration scheme to a logarithmic scheme. The importance of such a comparison is discussed in the following section where effect of image content on quality is made clear. In Figure 15, the CSF and the MTF from Figure 13 are multiplied by a hypothetical Gaussian Display MTF. Figure 16 is the same as Figure 15 except that the x-axis is a logarithmic scale.

In Figures 12 through 15, image quality is proportional to the area beneath the curves. The primary interest is in where (what spatial frequency ranges) most of the area below the curves are with respect to spatial frequency. From these graphs, one can generally infer that image quality is comprised mostly of low spatial frequency information. The finding is more robust when using the MTF of the eye as opposed to the CSF and when using a logarithmic integration scheme as opposed to a normal integration-across-spatial frequency (Figs. 14 and 16 versus Figs. 13 and 15). The effect of changing the integration scale from normal to logarithmic is to force the relationship between image quality and spatial frequency to follow a Weber-like function. For sound in the auditory sense and luminance in the visual sense, the psychological perception of the physical intensity of these variables is linearly related to the logarithm of these physical quantities. For the present work, this assumption has been applied to image quality and spatial frequency. This idea is mentioned again in the section "Image Quality Metrics and the Use of Image, Display, and Observer Characteristics," where the logarithmic metrics are more thoroughly discussed.

CHARACTERISTICS OF TWO-DIMENSIONAL STATIC IMAGES

When a static three-dimensional world is projected onto a two-dimensional space, depth information is lost or transformed through the projection. Although the perception of depth in two dimensions

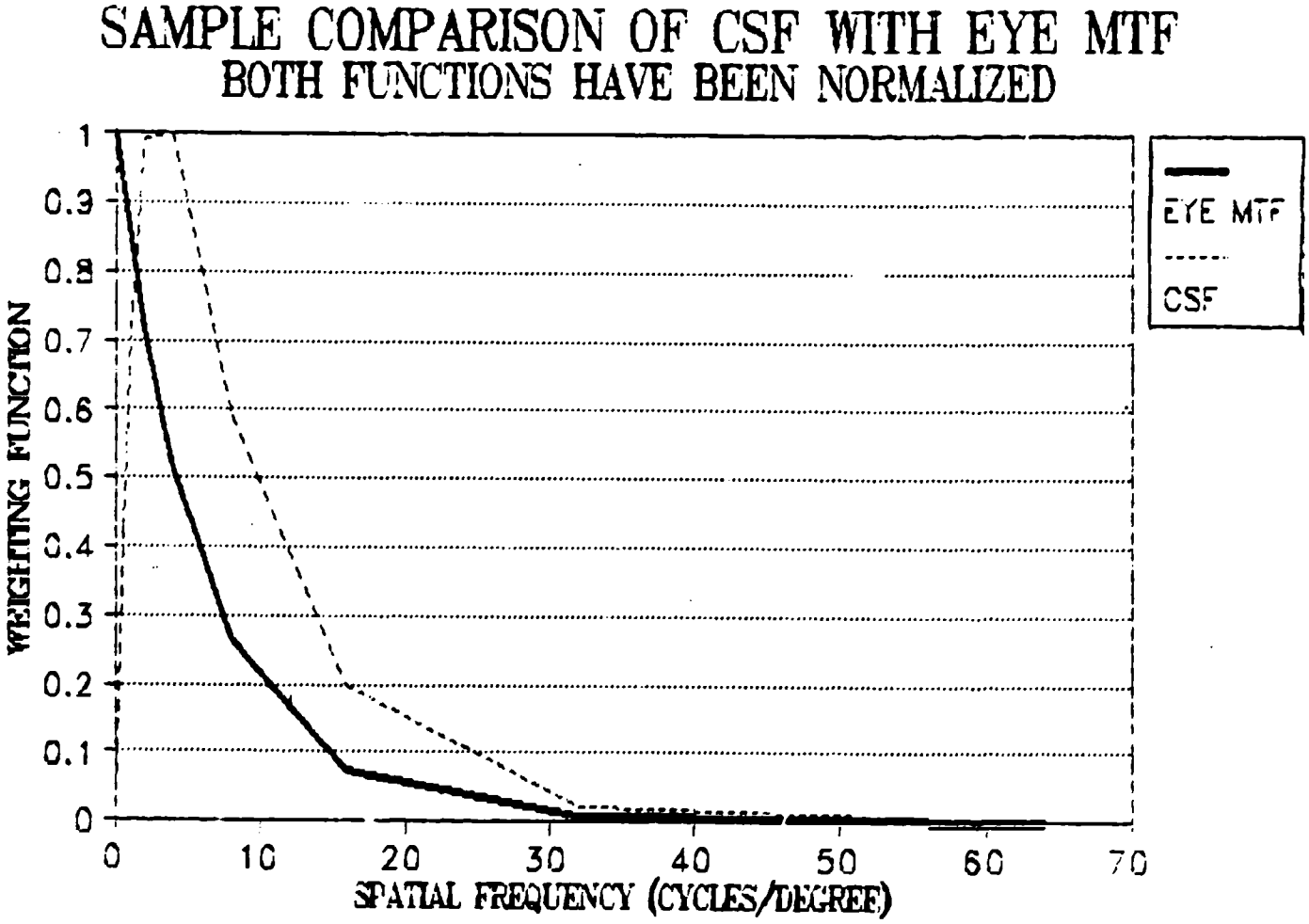


Figure 13
CSF versus MTF Weighting Function: Normal Plot

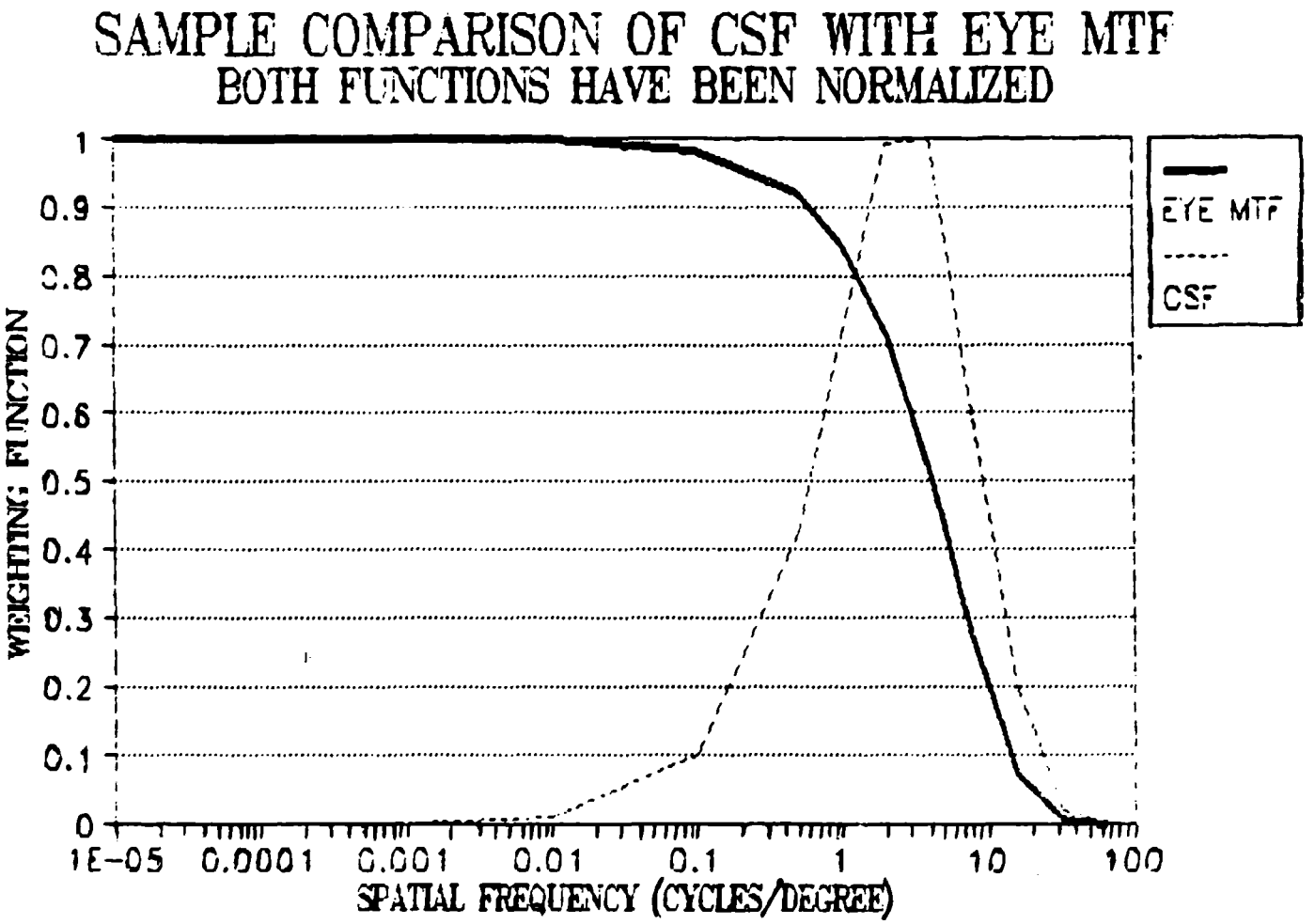


Figure 14
CSF Versus MTF Weighting Function: Logarithmic Plot

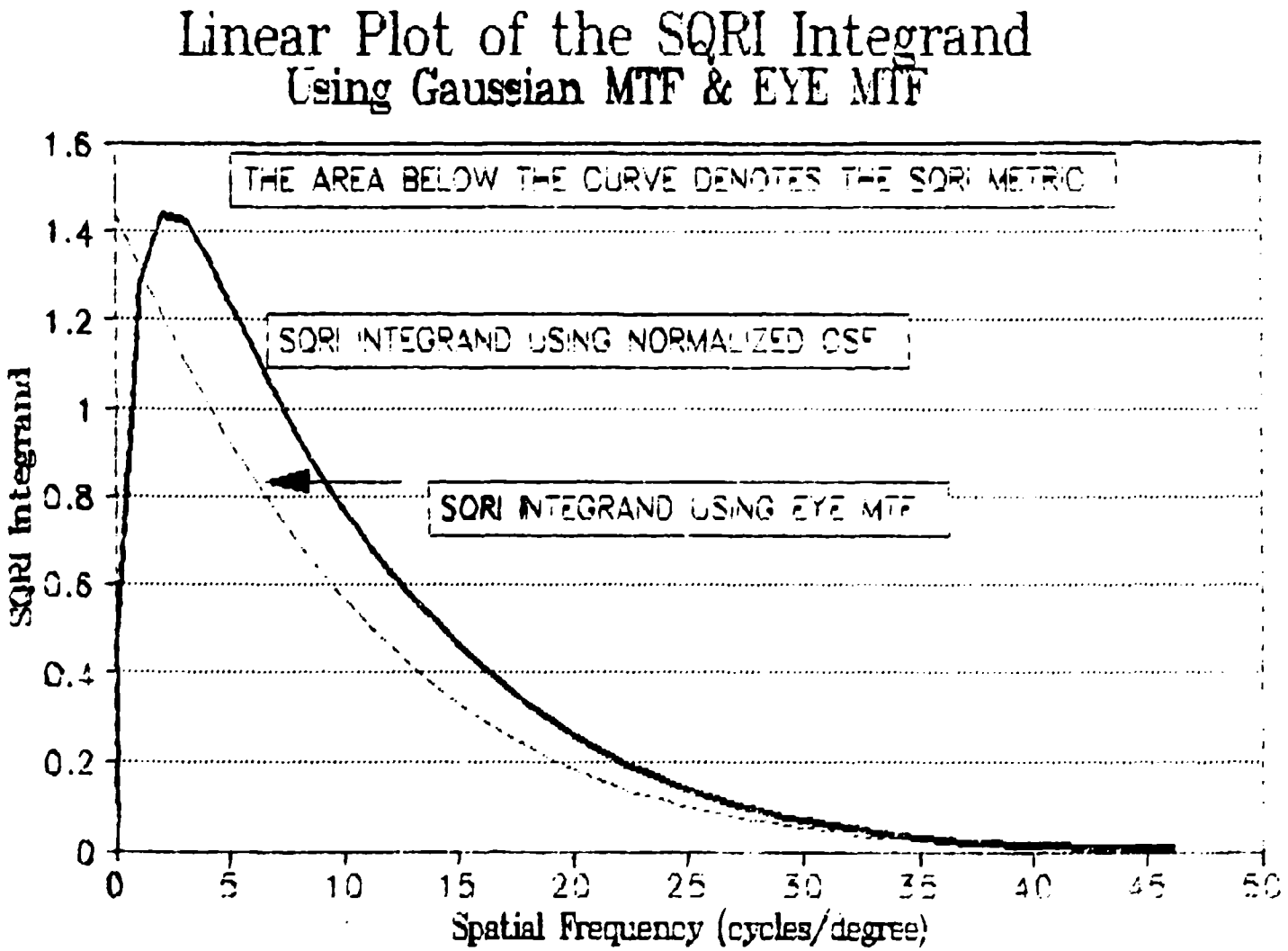


Figure 15
Normal Plot of the SQRI Integrand:
CSF versus MTF Comparison

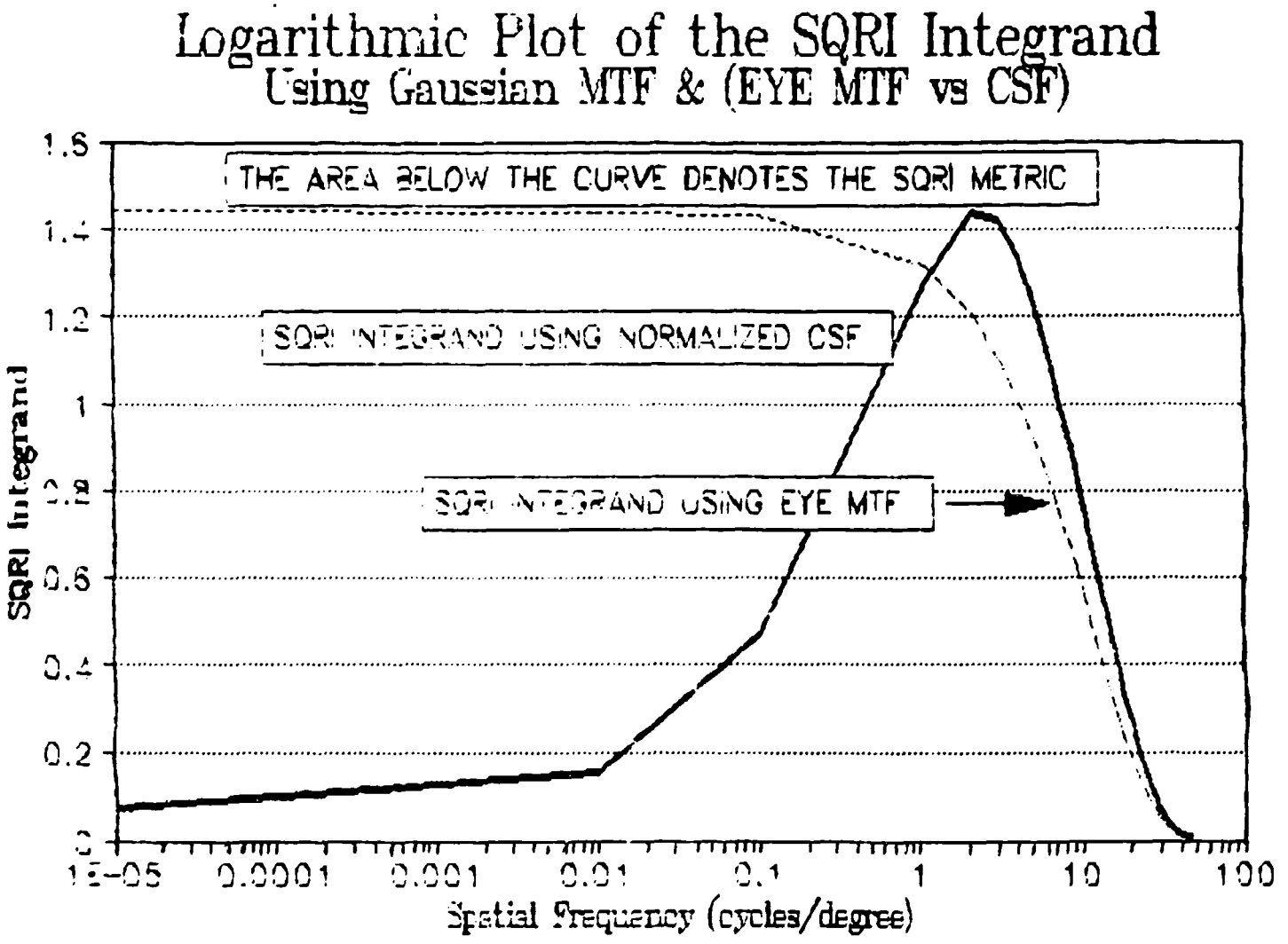


Figure 16
Logarithmic Plot of the SQRI Integrand:
CSF Versus MTF Comparison

is a key issue in simulator displays currently, it has received no attention, as yet, from an image quality perspective. In image quality, the focus currently lies with the two-dimensional display and its luminance, contrast, and resolution characteristics. This distinction is made here because it is clear that, at some point in time, image quality research will begin to concern itself with the perceptual aspects of computer-generated images and how well displays relate the appropriate visual cues to the viewer.

Spectral characteristics, or color in the psychological domain, contribute to perceived brightness, contrast, and resolution. At this point in time, however, the spectral components are weighted and summed according to their photometric or receptor (cone) weighting components and treated as a single scalar quantity (i.e., luminance in fL or cd/m^2).

Given scalar quantities (i.e., luminance) as points in a two-dimensional space, the static image may now be analyzed using a traditional Fourier approach. With digitized images, the discrete two-dimensional Fourier transform is applied to the image. The following section reports on the use of discrete Fourier transforms to characterize the luminance-varying content of imagery.

Global Versus Local Aspects of Images

Field (1987) and Hultgren (1990) reported on findings from the digitization and discrete Fourier analysis of natural images. Some of their findings are reproduced in Figures 17a and 17b. The results show that an overwhelming majority of the energy in the image is located at low spatial frequencies. Researchers in the image quality area have used such findings to suggest that image quality metrics should be weighted accordingly. That is, if images are overwhelmingly composed of low spatial frequency information, the ability of a display device to maintain image quality (e.g., contrast) at low spatial frequencies should be weighted more heavily relative to high spatial frequency capabilities.

Figures 18 through 22 portray static images used by Kleiss and Hubbard (1991) in the study of visual features important to low-level flight. These five images represent extremes from a multi-dimensional space obtained by Kleiss when pilots were asked to rate the similarities between a number of static images.

Figure 23 shows the magnitudes in a one-dimensional (the horizontal direction or along the raster lines) fast Fourier transform (FFT) of the five test images after the images were digitized into 512 X 512 elements. The results parallel those of Field (1987) and Hultgren (1990) in that the images are overwhelmingly composed of low spatial frequency information. In addition, note the difficulty in discriminating between the images

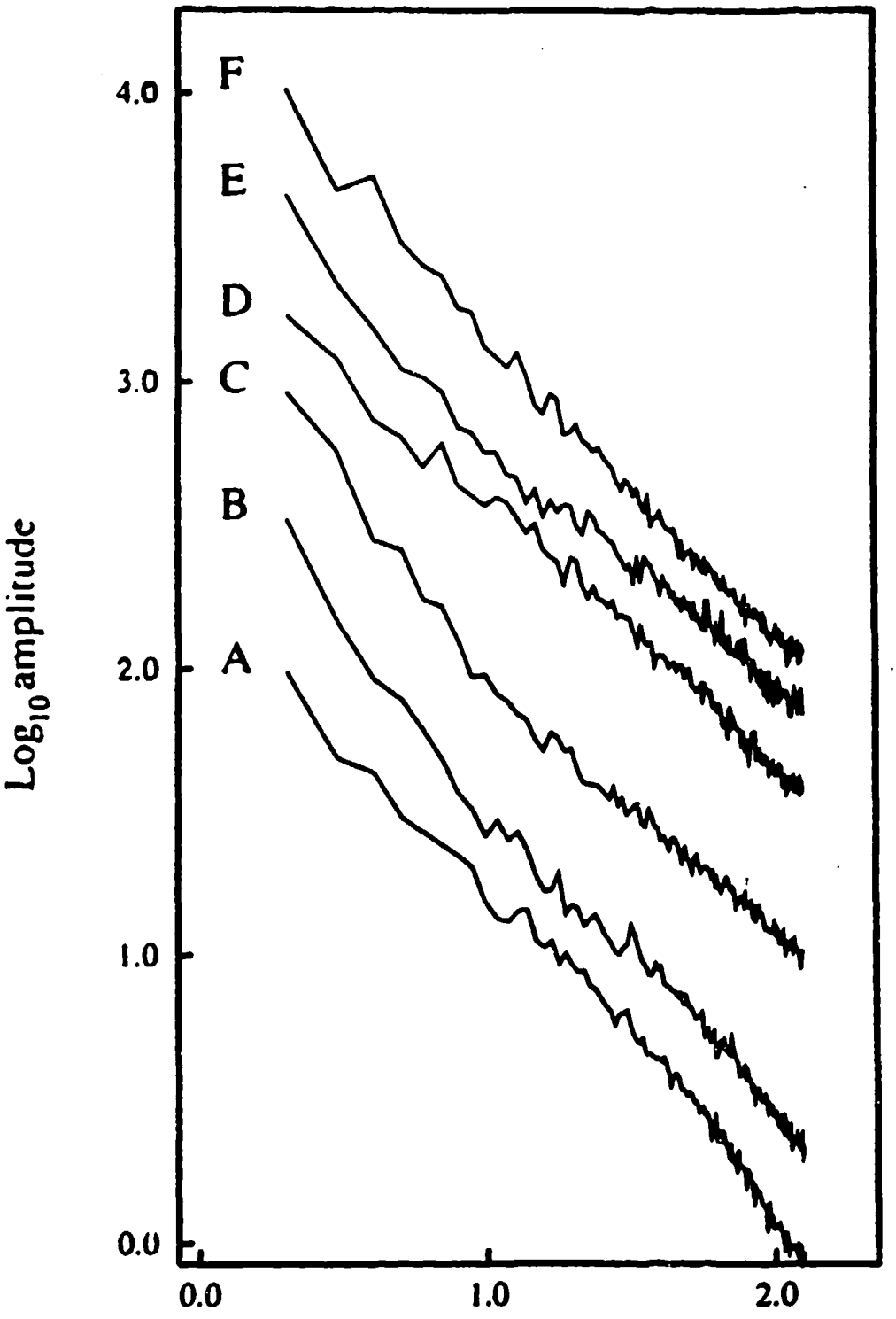


Figure 17a
Spatial Frequency Diagram from Field (1987). [Reproduced by permission from Optical Society of America, Vol. 4(12) 2379-2394].

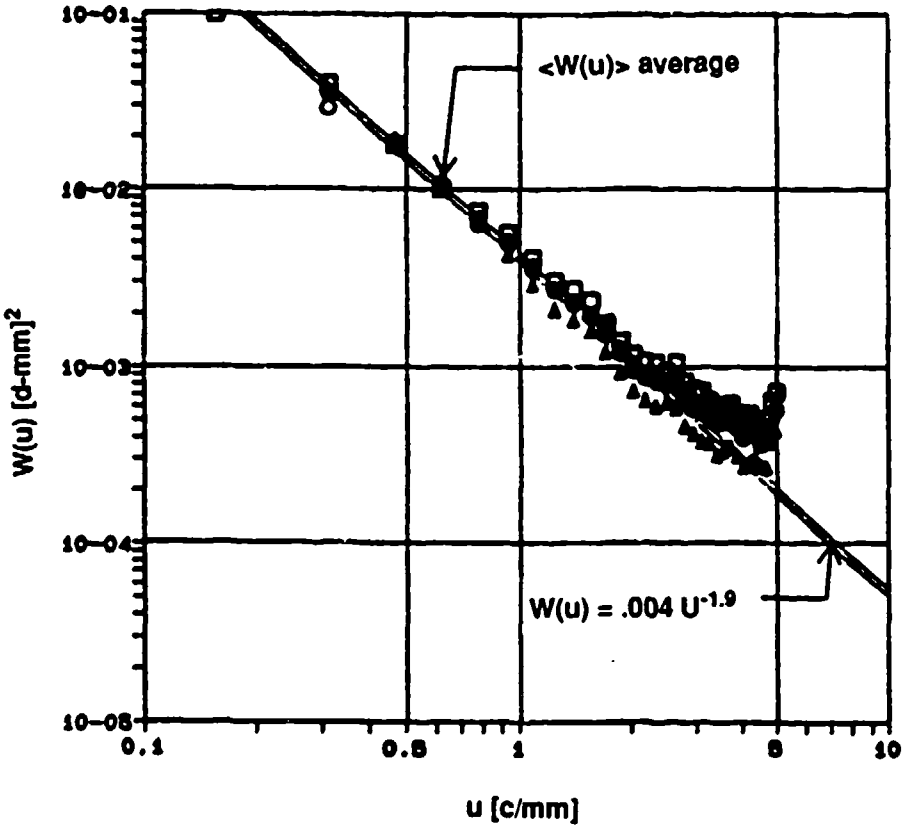


Figure 17b
Spatial Frequency Diagram from Hultgren (1990). (Reproduced with permission from SPIE, Human Vision and Electronic Imaging: Models, Methods, and Applications, Vol. 1249, 12-22)

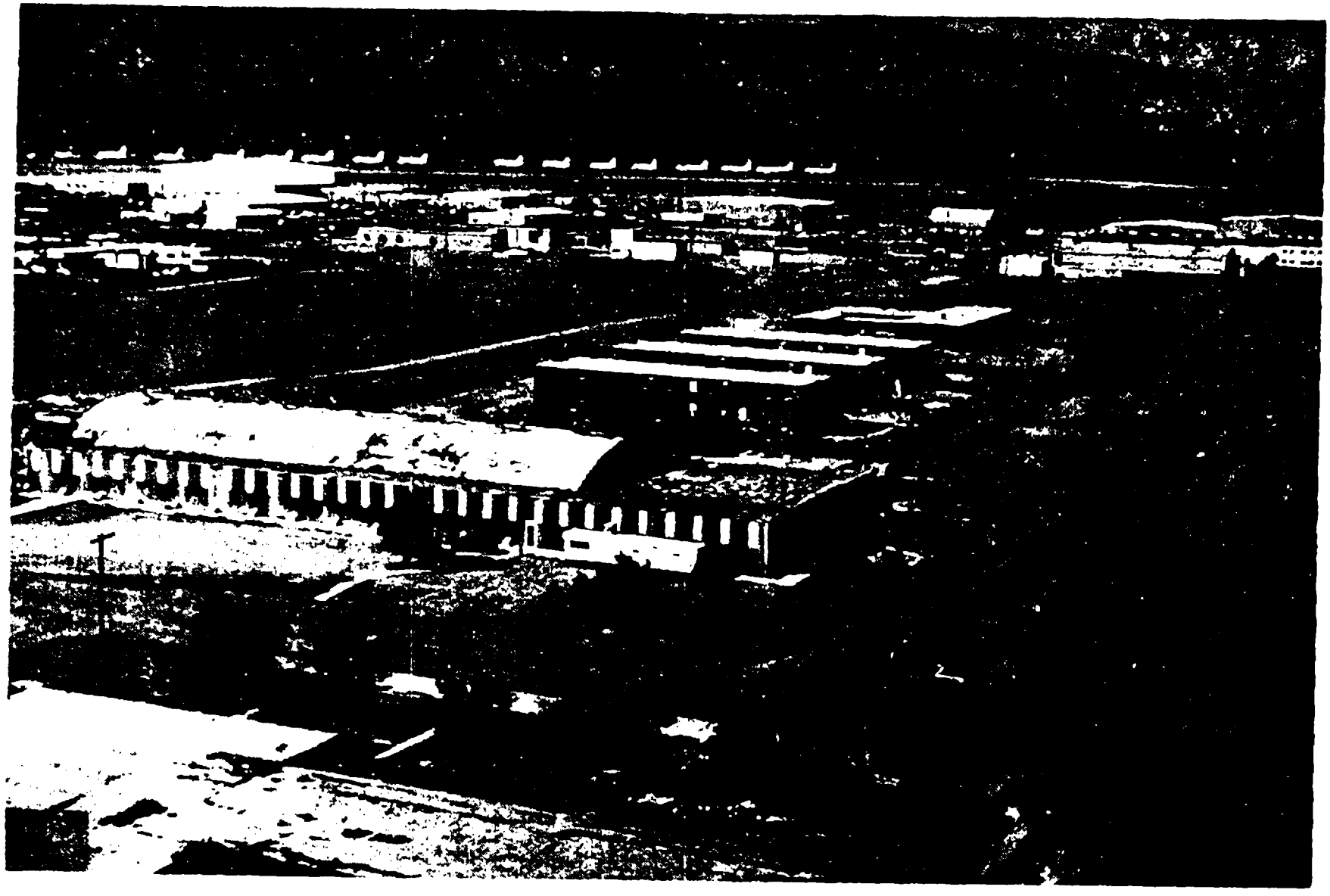


Figure 18
Low-Level Airfield Image from Kleiss (1991) MDS Study

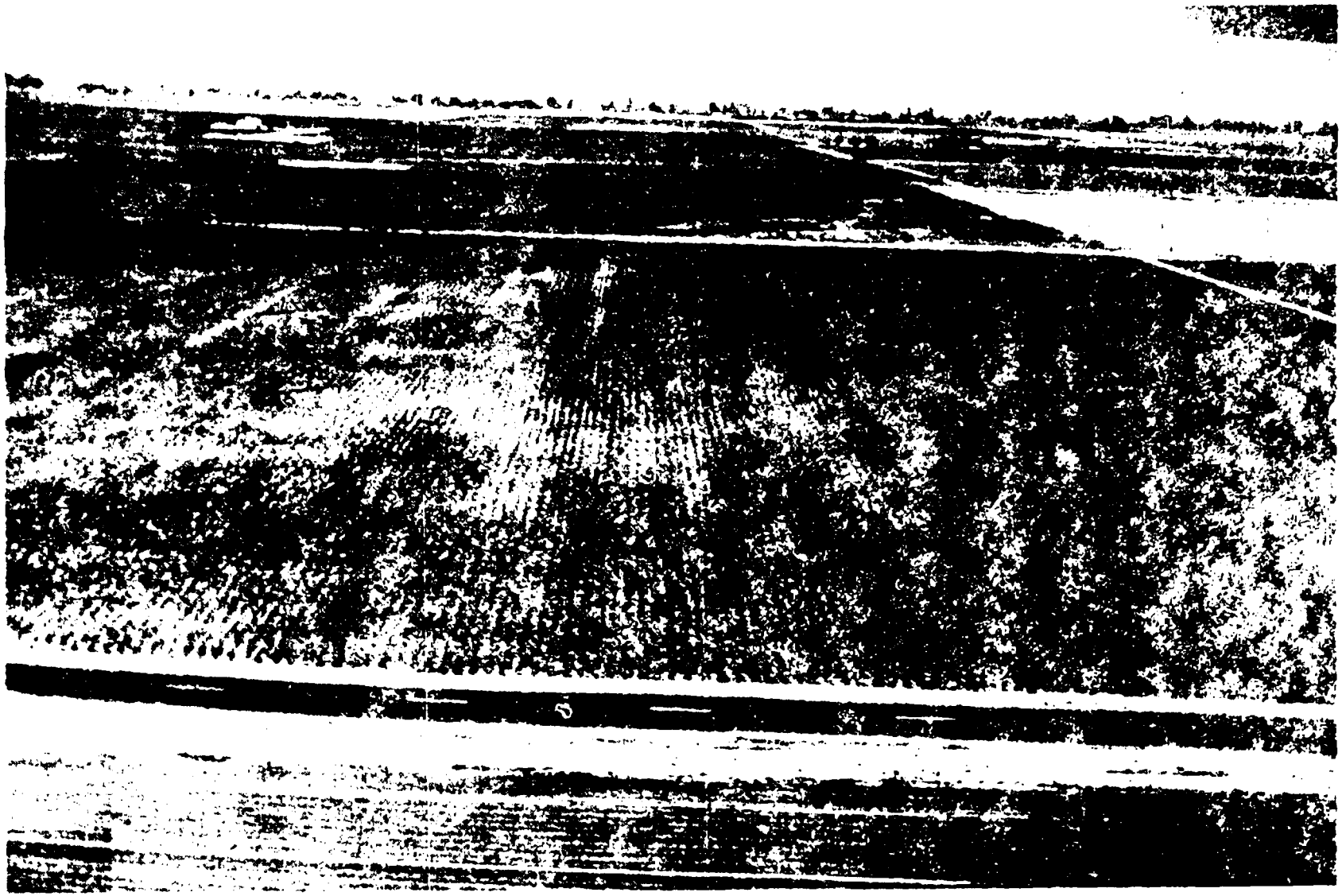


Figure 19
Low-Level Crop Image from Kleiss (1981) MDS Study

Figure 20
Low-Level Mountain Image from Kleiss (1991) MDS Study

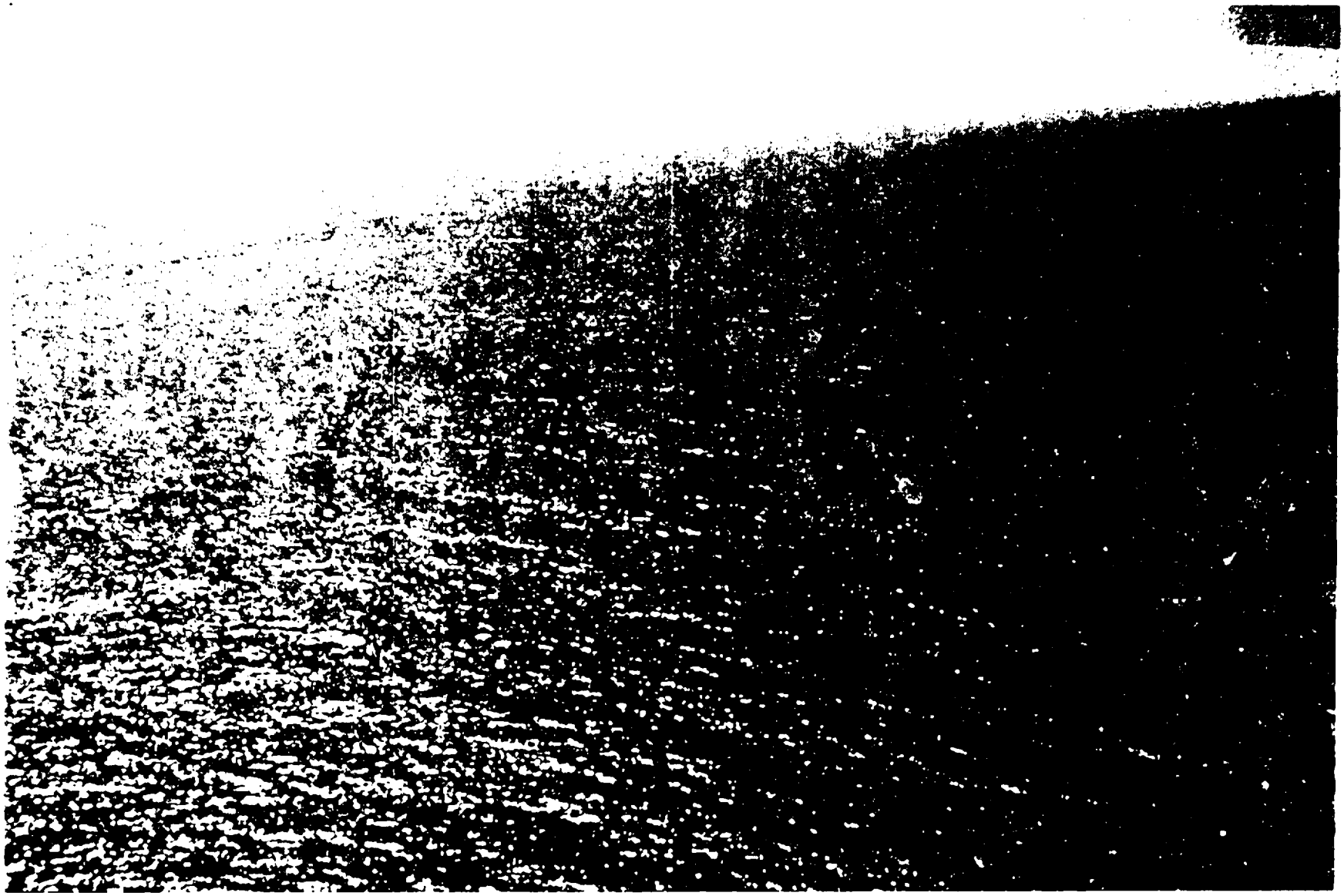


Figure 21
Low-Level Ocean Image from Kleiss (1991) MDS Study

Figure 22
Low-Level Pine Tree Image from Kleiss (1991) MDS Study

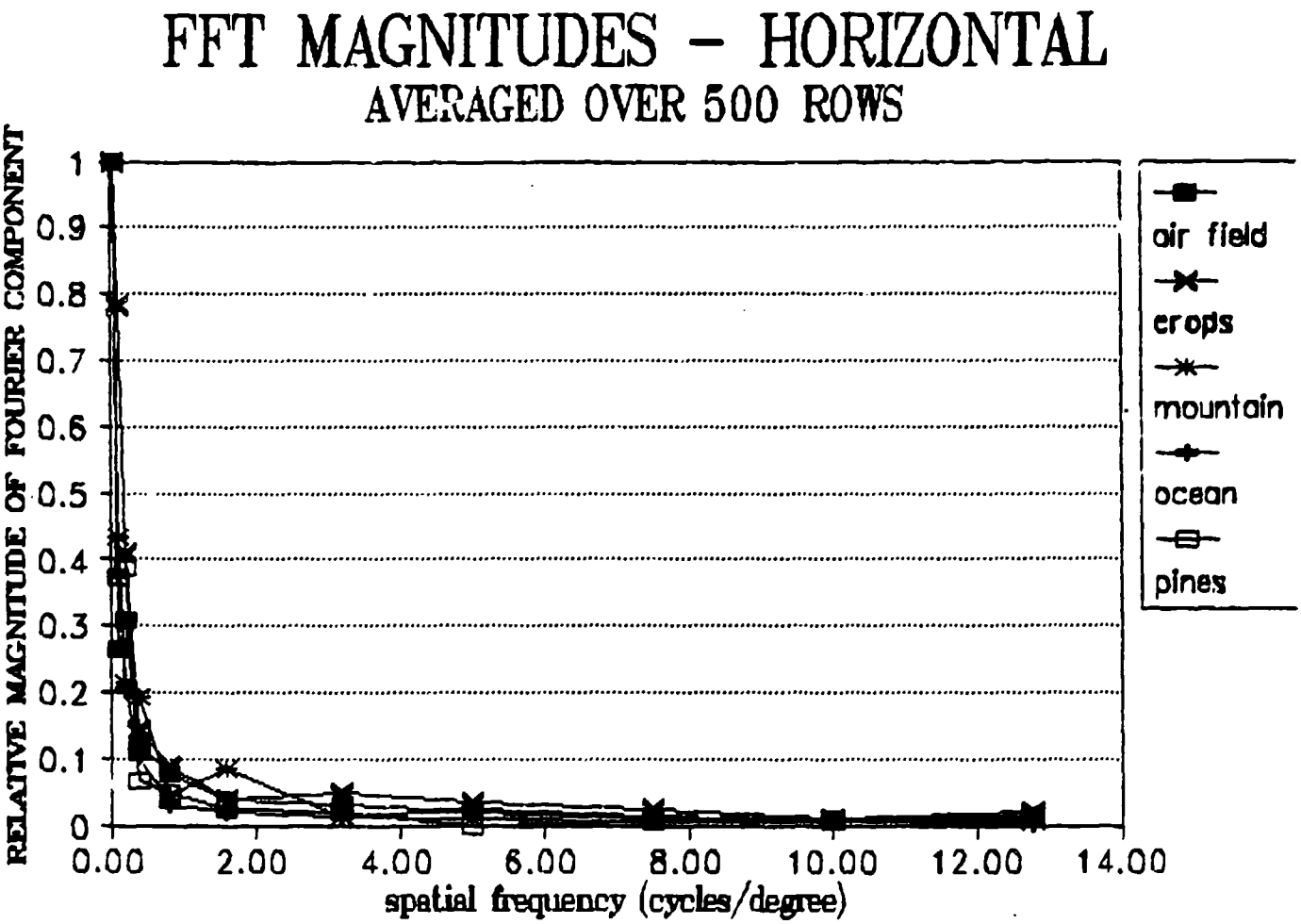


Figure 23
One-Dimensional Fast Fourier Transform of Kleiss Imagery

in the frequency domain. In spite of the distinctiveness of these images, the spatial frequency content of any one of the images could be used as a predictor of the other images. Field (1989) emphasizes this point using an example where random smudges in a two-dimensional image are shown to have the same power spectrum as a clear outline of a human face.

Figures 24 and 25 represent more localized Fourier analyses of the same five images. These results are based upon 32 point FFTs (Fig. 24) and 16 point FFTs (Fig. 25) taken at random locations in the images. Note the change in the relative amount of energy at higher spatial frequencies across Figures 23, 24, or 25 as the analysis moves from a global to a more localized basis.

The FFT of the complete image involves an averaging process of frequencies over the entire image. Most images contain large homogeneous blobs (in luminance content) which dominate the spatial frequency analysis. High spatial frequency information is contained in localized areas of the image, most likely areas on which observers tend to fixate during visual inspection of the image. In the next section, the implication of the distribution of energy in natural imagery for image quality metrics is discussed in more detail.

Image Characteristics and Their Relevance to Image Quality Metrics

The findings shown in Figures 23 through 25 are relatively robust. Static images in the everyday world are predominantly composed of low spatial frequency information. Taken only from the perspective of the image, then, and much like what any sampling theorem from descriptive statistics would tell us to do, it seems only natural to heavily emphasize these low spatial frequencies in image quality metrics.

Other image quality metrics of recent vintage (e.g., Barten, 1987; Granger & Cupery, 1972) have implicitly or explicitly taken this finding into account by integrating over the spatial frequency axis as a function of the logarithm of spatial frequency. Figure 26 provides an example of how integration by dln(u), the natural log of spatial frequency, affects integration relative to integration over du, linear spatial frequency, for the DART and LFOV displays in Figure 2. The area under both MTFs represents the value of the integration. In Figure 26, more than 90% of the area is below 1 cycle/degree of visual angle. In Figure 2, just the opposite is true. In addition, note that if the area under the curves is used as a metric, the LFOV is notably superior to the DART in Figure 26, while in Figure 2, the areas under the curve are more nearly equal, denoting equality of image quality.

The rationale to weight low spatial frequencies heavily based solely on the overall dominance of the energy at low frequencies in images (or a Weber's Law JND viewpoint) neglects our knowledge of

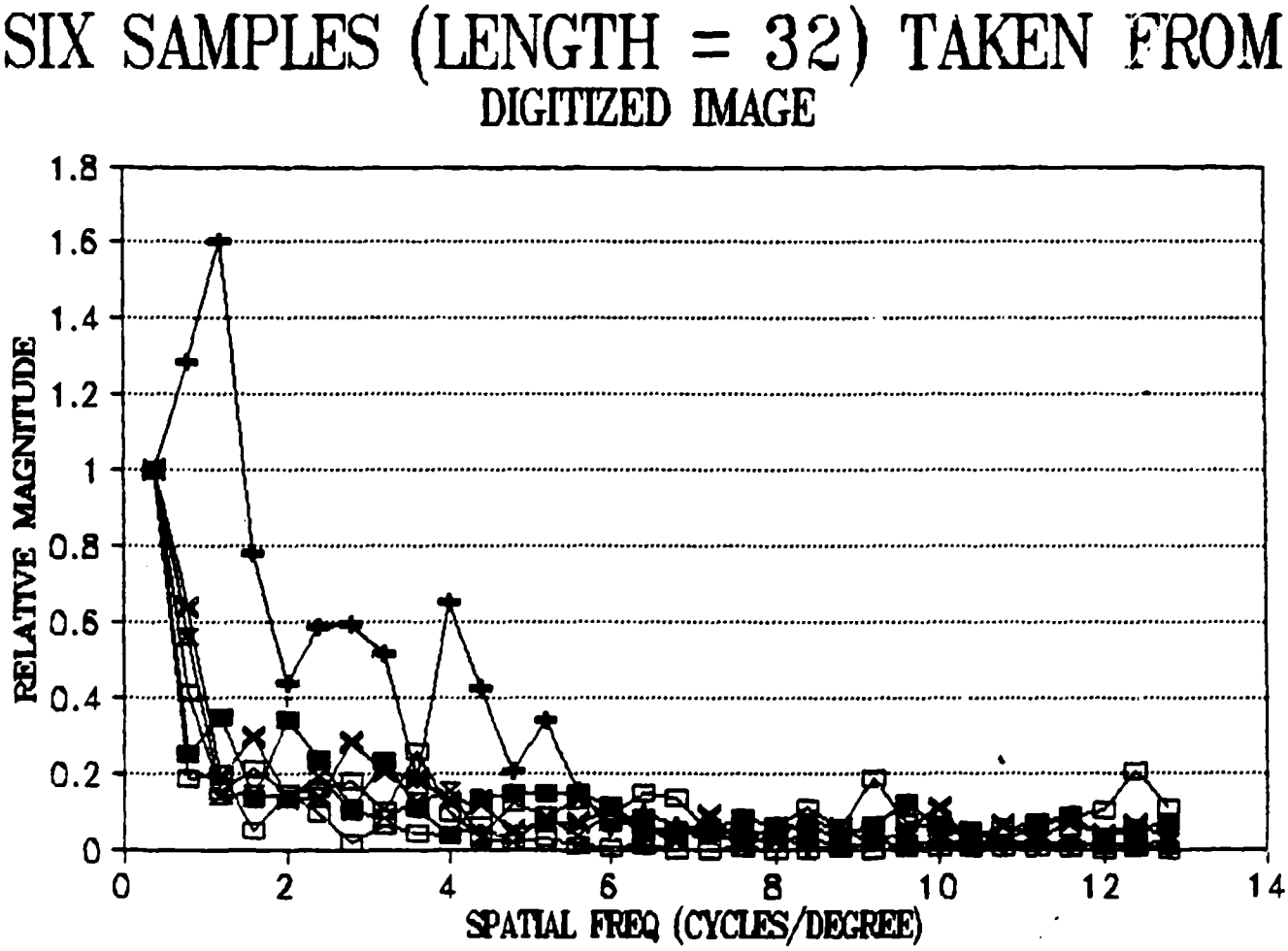


Figure 24
Localized 32-Point Fast Fourier Transform Example

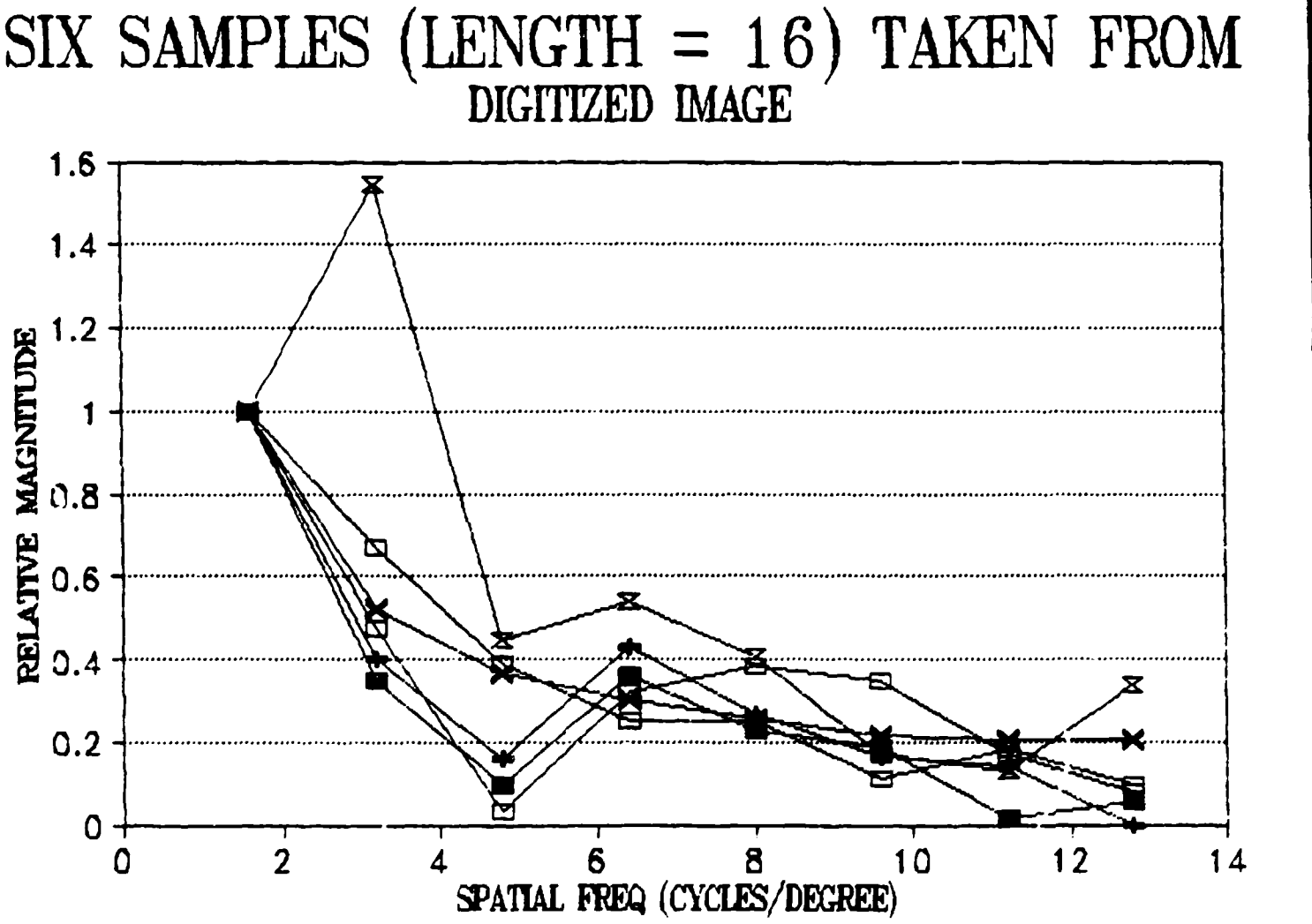


Figure 25
Localized 16-Point Fast Fourier Transform Example

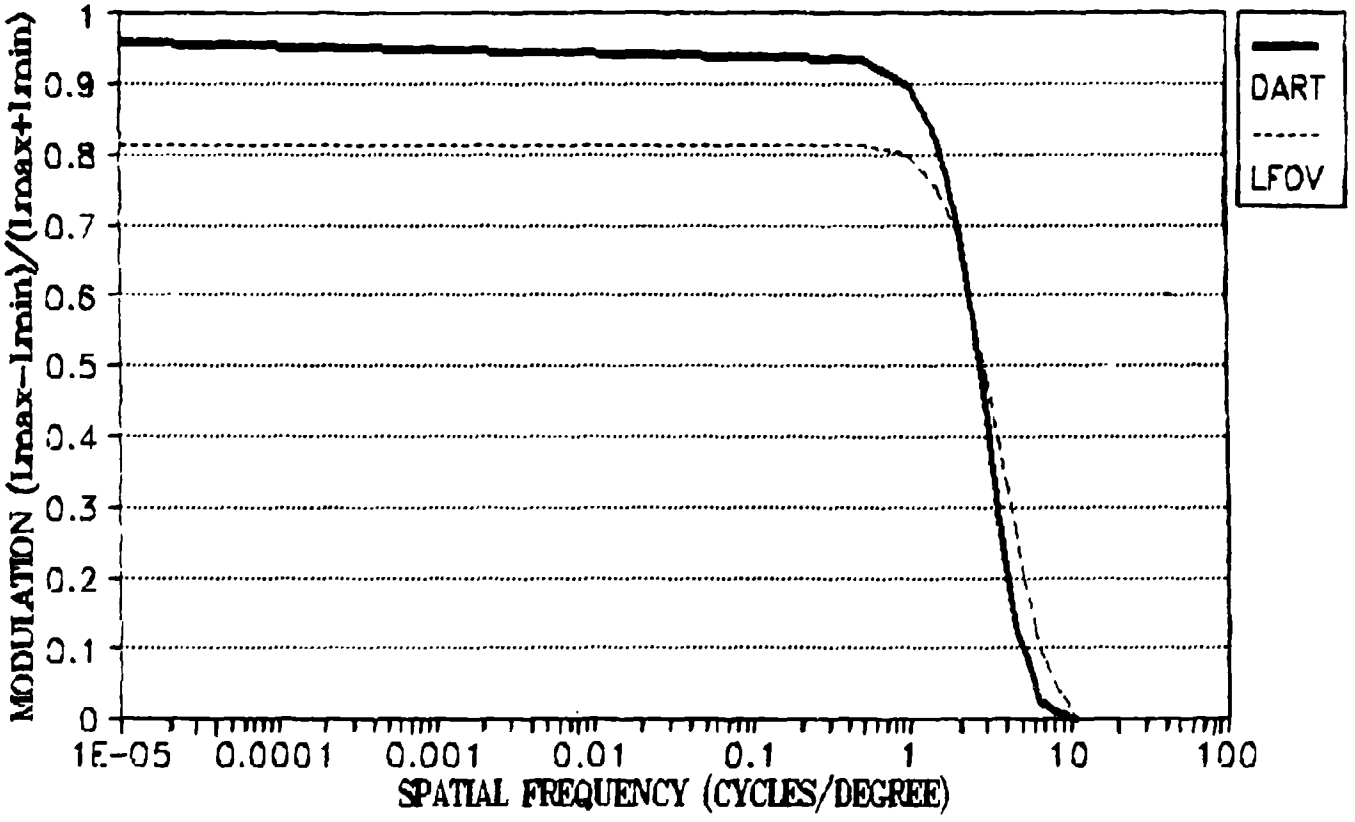


Figure 26
DART Versus LFOV MTF Comparison Plotted on Logarithmic Scale

the human visual system. Humans tend to foveate on high spatial frequency information and a disproportionate amount of the time in ocular excursions is spent visiting localities with high spatial frequency content. Of course, if one argues that our systems approach to image quality should combine the effect of each component (image, display, observer) separately, then the argument may be made that the logarithmic integration should be compensated for in the final stage of filtering. Neither the CSF nor the MTF of the eye attest to our propensity to foveate on a localized area of high spatial frequency. Inclusion for this propensity in an image quality metric would likely require a probability density function representing the relative amount of time spent fixating on localized areas within an image.

Without such compensation in the image quality metric, we can estimate what the effects will be on an image quality metric. Figures 26 and 27 are examples which drive home this point. Although device #2 retains more contrast over most of the spatial frequency range as shown in Figure 27, a plot of modulation by the natural log of spatial frequency (Fig. 28) will reverse the amount of area under the two curves. Any improvement in the high spatial frequency components of a visual display will easily be outweighed by much less compensation at low spatial frequencies.

One method of testing the relative contributions of high, mid, and low spatial frequency information in images would be to systematically filter out spatial frequency components of imagery. A variety of methods (scaling, paired comparisons) could then be used to empirically determine the quality of the imagery. However, modulating contrast within bands of frequencies while holding luminance constant is a complex task. This problem is discussed in more detail in the section entitled "An Experimental Approach for Examining the Effect of Display MTF on Perceived Image Quality" within this report.

Now that the components of the system (image, display, observer) have been discussed more thoroughly, a general introduction to image quality metrics from the literature shall be presented. In the next section, the MTF, the MTFA, the SQRI, and the SQF are presented as examples of image quality metrics in order to show how factors from the image, display, and observer are specifically incorporated.

IMAGE QUALITY METRICS AND THE USE OF IMAGE, DISPLAY, AND OBSERVER CHARACTERISTICS

In the previous sections, important characteristics of the image, the display, and the observer involved in the transmission of visual information were introduced. In this section, the contribution from each of these components to actual image quality metrics is examined using examples from the literature.

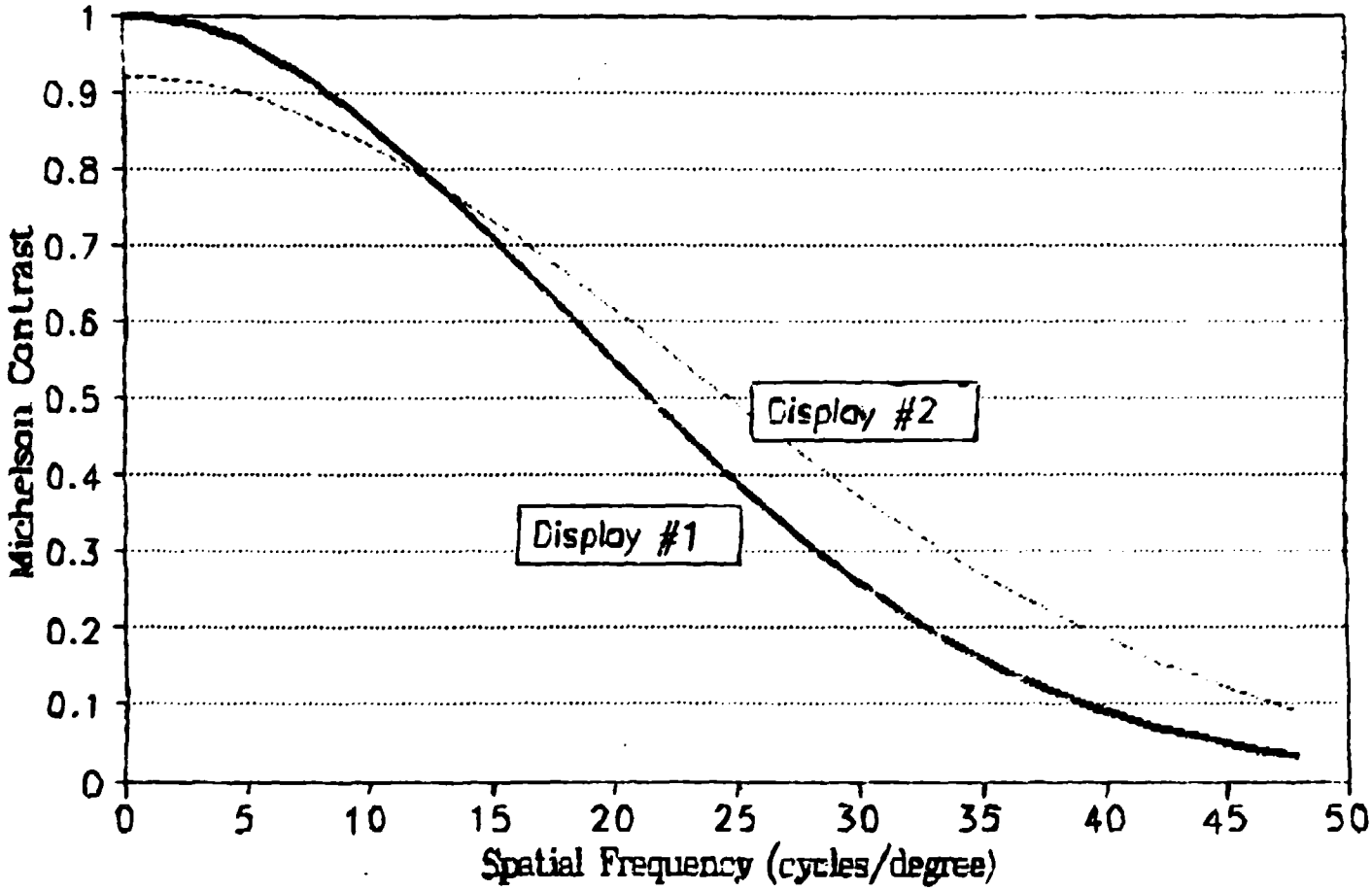


Figure 27
Comparison of Two Hypothetical Display MTFs on Linear Scale

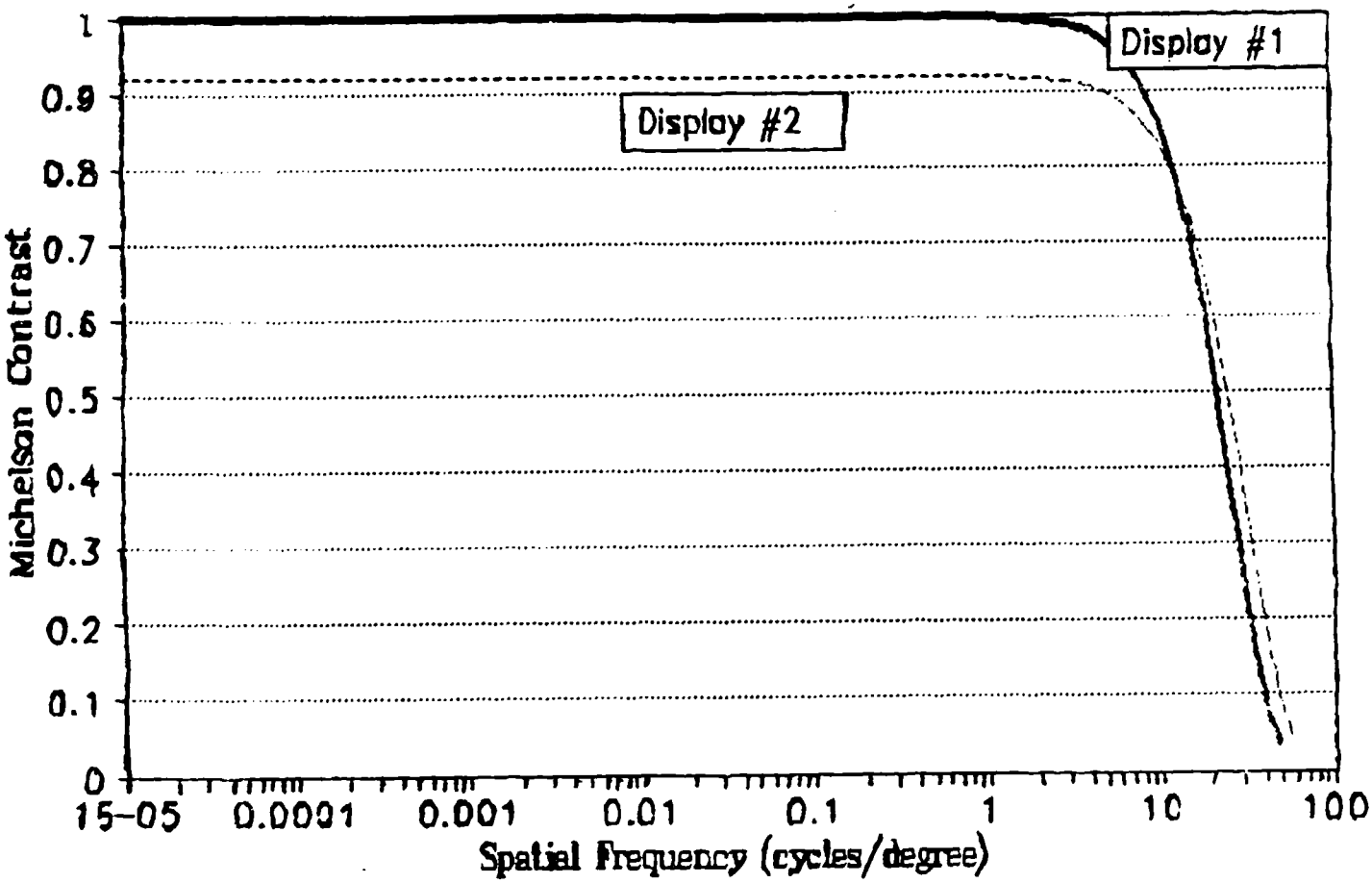


Figure 28
Comparison of Two Hypothetical Display MTFs on Logarithmic Scale

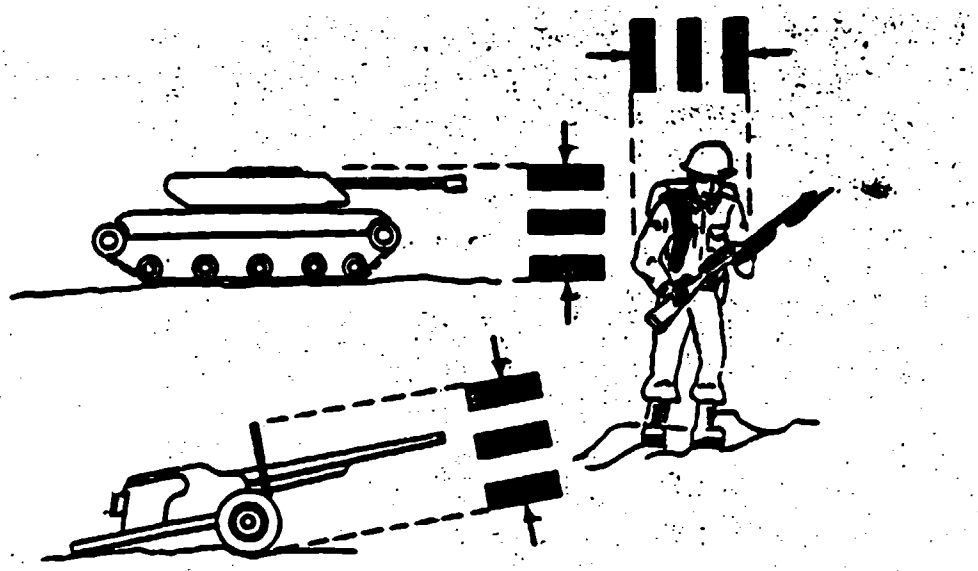
As suggested previously, the display MTF was the first factor used as an indicator of image quality. Strehl in 1902 (see Chapter 2 of Biberman, 1972) suggested using the area under the two-dimensional MTF as an indicator of image quality. In the 1940s, Schade (Biberman, 1972) is credited with popularizing the use of the display MTF as an indicator of image quality for televisions. As an image quality measure, he used the upper frequency cutoff of a square wave which had an equal amount of area under the curve as the display MTF.

Both of the techniques discussed are compact and efficient methods of obtaining a metric or measure of image quality. They disregard many parameters which are critical to image quality but were probably sufficient for comparing displays from their time period. More recently, though, technological changes have allowed for the existence of a large variety of visual displays based upon a variety of technologies. Owing to this variety, the resulting imagery differs along many dimensions and the different technologies will incur trade-offs across these dimensions.

In some cases, improvements in technology could permit the generation of imagery which surpassed the capabilities of the observer. The inclusion of the CSF function along with the display MTF served as an attempt to subtract from the display MTF information which could not be used by the observer. The MTFA (Snyder, 1985) serves as an example of such a metric.

As shown previously, the CSF is the inverse of the minimum amount of contrast required to discriminate a sinusoidal waveform from a homogeneous field of equal luminance. The inverse of the CSF may be designated as the Demand Modulation Curve (DMC) or that amount of contrast which the visual system demands for discrimination. Snyder (1985, Chapter 4) denotes this function as the Contrast Transfer Function (CTF). Scott (1966) developed a similar function for the resolvability of a three-bar pattern which he called the Demand Modulation Function (DMF). The use of a three-bar pattern is a derivation of the Johnson Criteria (Johnson, 1958 - see Fig. 29) where targeting performance is equated with the resolution of bar patterns. It is interesting to note that the spatial frequency approach in image quality (i.e., required modulation as a function of spatial frequency) is not that far removed from empirical work performed in the late 1950s.

Figure 30 shows a plot of a one-dimensional display MTF along with three demand modulation curves. The three DMCs were generated for three different levels of average display luminance using van Meeteren's approximation as presented in Equation 18. By computing the area under the display MTF in Figure 30, a simple metric is formed. Conceptually, however, it can be argued that any display modulation capabilities below the threshold of the visual system, as determined by the DMC should not contribute to a metric. In



METHOD OF OPTICAL IMAGE TRANSFORMATION

TARGET	RESOLUTION PER MINIMUM DIMENSION IN LINE PAIRS			
BROADSIDE VIEW	DETECTION	ORIENTATION	RECOGNITION	IDENTIFICATION
TRUCK	0.90	1.25	4.5	8.0
M-48 TANK	0.75	1.20	3.5	7.0
STALIN TANK	0.75	1.20	3.3	6.0
CENTURION TANK	0.75	1.20	3.5	6.0
HALF-TRACK	1.00	1.50	4.0	5.0
JEEP	1.20	1.50	4.5	5.5
COMMAND CAR	1.20	1.50	4.3	5.5
SOLDIER (STANDING)	1.50	1.80	3.8	8.0
105 HOWITZER	1.00	1.50	4.8	6.0
AVERAGE	1.0 ± 0.25	1.4 ± 0.35	4.0 ± 0.8	6.4 ± 1.5

Figure 29
Johnson Criteria for Targeting Performance. Analysis of Image Forming Systems. (Reproduced with permission from SPIE, Infrared Design, Image Intensifier Symposium, (1958), Vol. 513, Part I)

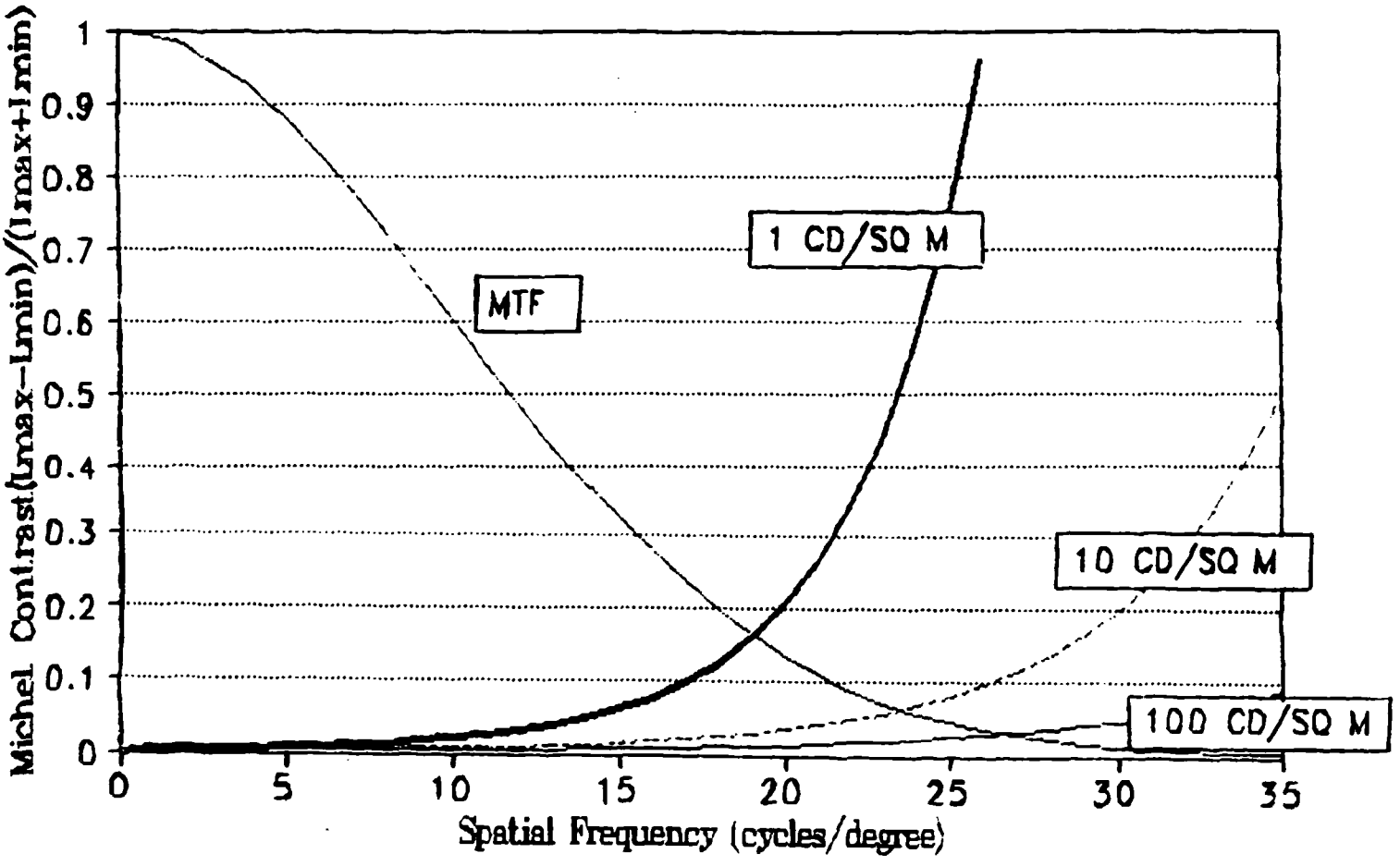


Figure 30
Display MTF and Three Demand Modulation Curves (DMCs)

Figure 30, this would translate into subtracting the area below the DMC from the area below the display MTF for all spatial frequencies. Mathematically, this idea is as follows:

$$MTFA = \int_{u=0}^{u=umax} [MTF(u) - DMC(u)]\, du \qquad (25)$$

where u is spatial frequency in, for example, cycles per degree of visual angle and umax represents the highest frequency displayed by the display device.

The important considerations of the MTFA are as follows. First, note that the integration across spatial frequency in Equation 25 is linear or that the integration is with respect to u, the spatial frequency. That is, all spatial frequencies are weighted equally in their contribution to the metric. This weighting is in contrast to the more-popular logarithmic weighting scheme discussed previously and shown in the metric examples to follow. The second distinctive attribute of the MTFA is that the MDC, or the inverse of the CSF, plays the part of a subtractive threshold in the integrand, not a multiplicative weight. This use is distinctively different from that of most approaches which employ the human filter as a multiplicative weight.

Display luminance may be incorporated indirectly as a parameter into the MTFA from the MDC. As display luminance increases, the demand modulation curve either remains the same or decreases for all spatial frequencies, depending upon the model employed for representing Contrast Sensitivity. For example, the van Meeteren Estimate (Equation 19) incorporates display luminance although other estimates may not. The end result is that the MTFA can be made to increase with improvements in display luminance. This ordinal implication, though, says nothing about the comparative effects of display luminance versus display MTF on image quality.

From the MTFA, we proceed to discussion of metrics which employ the human filtering as a multiplicative weight. Granger and Cupery (1972) presented the Subjective Quality Factor (SQF) as an objective figure of merit for testing MTF shapes. The SQF is given as:

$$SQF = K \int_{u=10}^{u=40} MTF(u)\; d(\ln(u)). \qquad (26)$$

u in Equation 26 above denotes spatial frequency in cycles per millimeter at the retina. The integration in Equation 26 is with

respect to the natural log of u. The premultiplier K is a normalizing constant such that:

$$K = \frac{1}{\int_{u=10}^{u=40} 1 \, d(\ln(u))} \tag{27}$$

which causes the SQF to have lower and upper bounds of zero and 1, respectively.

Granger and Cupery (1972) used a logarithmic integration over the spatial frequency axis under the assumption that image quality, like other psychophysical sensations, should follow a Weber's function. That is, psychophysical sensation (e.g., perception of brightness, loudness) as measured in JNDs from relative observer thresholds, is a logarithmic function of the magnitude of the physical phenomenon.

Another interesting point of the SQF is that the integration has lower and upper bounds of 10 and 40 cycles per millimeter on the retina. These limits were used for integration based upon the finding that the eye is most sensitive within these limits (Schade, 1964). If we assume that the focal length from the lens of the eye to the retina is approximately 22.42 millimeters (Pugh, 1988), 10 and 40 cycles/millimeter translates into 3.91 and 15.65 cycles/degree of visual angle, respectively. In an indirect fashion, this assumption provided a bandpass filter for the human visual system which was identically 1 between 3.91 and 15.65 cycles/degree, and zero outside this range.

Barten (1987, 1989, 1990) introduced the SQRI measure. This measure uses the CSF as a weighting function for the display MTF. The mathematical form of the SQRI is as follows:

$$SQRI = \frac{1}{\ln(2)} \int_{u=o}^{u=umax} [MTF(u)\ CSF(u)]^{\frac{1}{2}} \, d\ln(u) . \tag{28}$$

where u is spatial frequency in, for example, cycles per degree of visual angle, CSF denotes the Contrast Sensitivity Function, and MTF denotes the Modulation Transfer Function or, equivalently, the modulation depth curve. The premultiplier of 1/ln(2) was chosen so that if the MTF was equivalent to the inverse of the CSF over a 1-log unit range and zero elsewhere, Equation 28 would integrate to a value of 1.

The unit of measurement for Equation 28 is JNDs. A perceptual JND is operationally defined as a 75 correct response rate in a two-alternative, forced choice experiment. Hypothetically, then, if two images, A and B, were presented and image A had an SQRI value which was one JND higher than the image B, the observer would

prefer image A on 75% of the trials and image B on 25% of the trials.

The SQRI in Equation 28 and the SQF in Equation 26 use a logarithmic integration over the spatial frequency axis. Figures 31 and 32 are plots of the integrand in Equation 28 (the SQRI) with respect to the logarithm of spatial frequency and spatial frequency, respectively. In Figures 31 and 32, Equation 19 was used to approximate the CSF, and the rightmost equality of Equation 16, a Gaussian profile, was used to approximate the display MTF. The area under the curve in Figure 31 represents the value of the SQRI in Equation 28 for the sample CSF and MTF employed. The area under the curve in Figure 32 represents the value of the SQRI if a linear integration was performed in Equation 28.

By comparing the curves in Figures 31 and 32, the information content of the metric becomes evident. In Figure 31, more than 90% of the integral is derived from spatial frequencies less than 10 cycles/degree. The conclusion is that a logarithmic integration tends to weight low spatial frequency information (i.e., modulation capability) very heavily in a display device relative to high spatial frequency information. Referring back to the discussion on global versus local aspects of images, this weighting is consistent with the spatial frequency content of natural imagery. As mentioned, though, observers tend to spend a majority of time fixating or foveating on high spatial frequency information in images. For image quality purposes, this fact may tend to outweigh the predominance of low spatial frequency information within images.

In order to test such hypotheses (i.e., the relative importance of information within specific bands of spatial frequency), it is necessary to present stimuli which have been differentially filtered in the spatial frequency domain. This is a complex task, though, because of the difficulty in filtering specific spatial frequencies and, at the same time, maintaining constant energy or luminance in an image. This problem is discussed further in the next section.

As an example of empirical tests of the metrics described here, many of the referenced papers used the following procedures. First, observers rank order or use a Likert Scale rating to show their preference for images from the experimental displays. Next, the appropriate image quality metric is computed for each experimental display. The rank orders or ratings are then correlated with the computed metric for the display of interest. Many authors plot the empirical rating on the x-axis and the metric prediction on the y-axis and test for a linear fit (i.e., a correlation). After collecting the empirical ratings, it is then a straightforward matter to compare correlations obtained by different metrics in order to determine which metrics perform best.

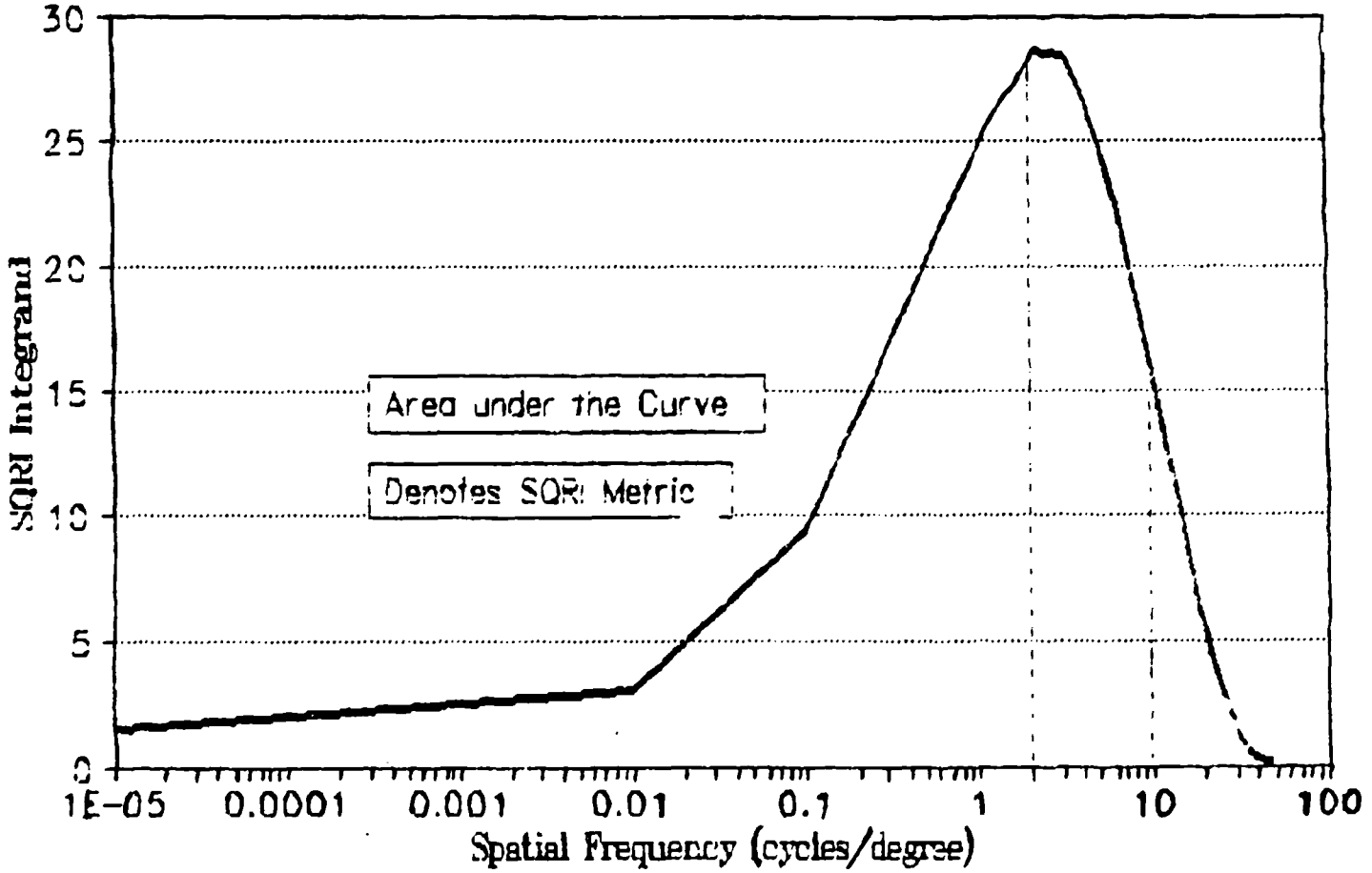


Figure 31
Logarithmic Plot of the SQRI Integrand

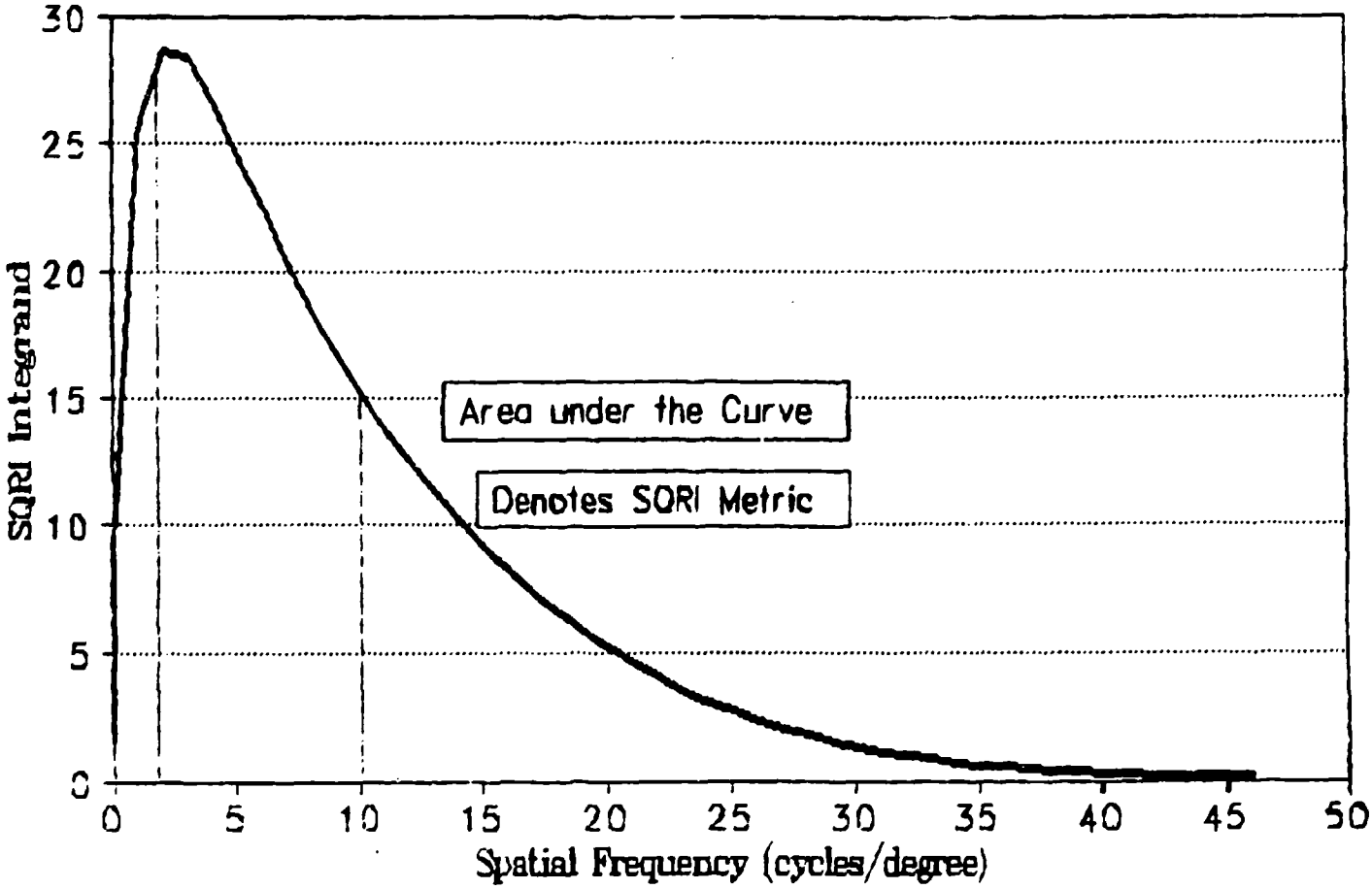


Figure 32
Linear Plot of the SQRI Integrand

This correlational method is simple and straightforward. Its drawback is that it is very insensitive to changes in many display parameters and lacks power in testing many of the multidimensional aspects of interest in image quality. For example, a linear fit to the normal ogive (the cumulative normal distribution curve) predicts 97.7% of the variance (r_{xy} > .99) for the ogive in a 95% confidence interval around the mean. Processes which are normally distributed can be almost perfectly approximated by triangular distributions or linear fits. The analogy in image quality is that the metrics are correlated with empirical measures of image quality. Whether they capture the essence of factors contributing to the process, though, is unclear.

In the remaining two sections of this report, an experimental methodology is described for manipulating individual display parameters in imagery and measuring changes in observer responses.

AN EXPERIMENTAL APPROACH FOR EXAMINING THE EFFECT OF DISPLAY MTF ON PERCEIVED IMAGE QUALITY

As shown throughout this report, the display MTF is a cornerstone in the construction of image quality metrics. For display devices, the MTF denotes the amount of Michelson Contrast available from a sinusoidal waveform as a function of spatial frequency. As pointed out, ambiguities exist in the empirical measurement and development of the display MTF. These ambiguities carry over into image quality metrics.

As shown earlier on measuring luminance modulation from display devices, the major ambiguity lies in the fact that a true MTF is identically 1 at DC or zero spatial frequency. In classical linear systems, this equivalence reflects the fact that signal energy is neither lost nor gained from input to output. The Michelson Contrast reaches a value of 1 only when the minimum luminance from the screen is identically zero. For most displays, two factors will contribute to the minimum luminance, ambient light and the dark or minimum luminance level of the display. Factors affecting ambient illumination in the immediate environment include not only external lighting within the environment but also the display (and its physical size) as light is reflected within the environment. When modulation depth curves are generated by the direct measurement method (see Beaton, 1989, or Kelly, 1992), the Michelson Contrast at zero frequency (i.e., the DC value on the MTF curve) is computed using the maximum and minimum or dark luminance values measured from the display. For practical displays, this modulation will never reach a value of 1.

It should be clear, then, that when applying empirical display MTFs or modulation depth curves in metrics, curves should be normalized to a value of 1 at zero frequency with the knowledge

that this is not a true estimate of display contrast in the actual environment. If one wishes to isolate the display from the surrounding environment and its ambient illumination, the ambient illumination may be subtracted from the minimum luminance measured from the display. This may yield modulation near unity at zero spatial frequency.

In this section, hypothetical display MTFs are generated which have a modulation depth of 1 at zero frequency. Our interest is in being able to simulate the effect a range of display MTFs will have on images and empirically measure observer responses. The visual and psychological literature has an abundance of data concerning observer sensitivity to sinusoids and square wave forms of varying contrast. There are little data available concerning modulation filtering of real-world images. The next subsection discusses in further detail the digital filtering of the experimental images.

Filtering, Display, and Observer Comparison of Experimental Images

In order to study the effect or how variations in the display MTF affect perceived image quality, a number of assumptions and simplifications were made. Gaussian MTFs were chosen to represent realistic display MTFs. As shown earlier, Gaussian MTFs can be efficiently represented and manipulated in both the spatial and the frequency domains. Five sample MTFs were chosen as filters. Frequency and spatial representations of these filters are shown in Figures 33 and 34 respectively. Figure 35 is a normalized version of Figure 33 where all values in the curve have been divided by the modulation at zero or DC frequency. All three figures (i.e., 33, 34, and 35) are one-dimensional representations for simplified viewing. The actual filters are two-dimensional with symmetry holding across dimensions (i.e., independence holds across the two dimensions). In Figure 33, each of the curves are of the form:

$$\begin{aligned} & MTF(f) = ae^{-bf^2} && (1\ dimension) \\ OR \quad & MTF(f_1, f_2) = ae^{-b(f_1^2+f_2^2)} && (2\ dimensions) \end{aligned} \tag{29}$$

where f_1 and f_2 denote spatial frequency in the vertical and horizontal dimensions measures (f denotes spatial frequency in a single direction) in cycles/degree of visual angle. Each curve may now be identified uniquely through the tuple (a,b) as shown in the legend of Figure 33. Applying the inverse Fourier transform, these curves can now be represented in the spatial domain (Figure 34) using the equation:

$$\begin{aligned} & h(x) = \frac{a\sqrt{\pi}}{\sqrt{b}} \frac{e^{-\pi^2 x^2}}{b} && (1\ dimension) \\ OR \quad & h(x_1, x_2) = a\frac{\sqrt{\pi}}{\sqrt{b}} e^{-\frac{\pi^2(x_1^2+x_2^2)}{b}} && (2\ dimensions). \end{aligned} \tag{30}$$

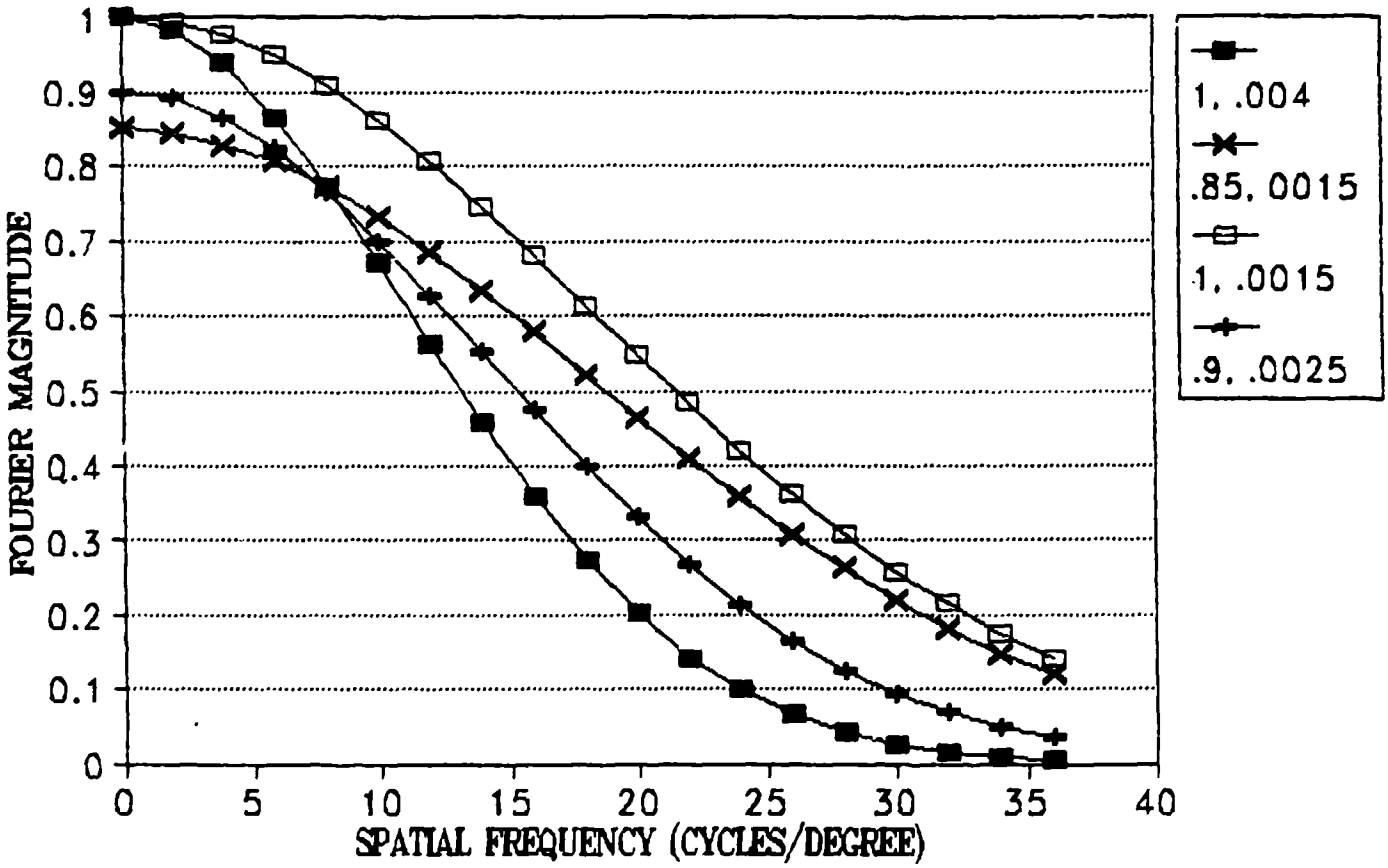


Figure 33
One-Dimensional MTFs for Experimental Filters

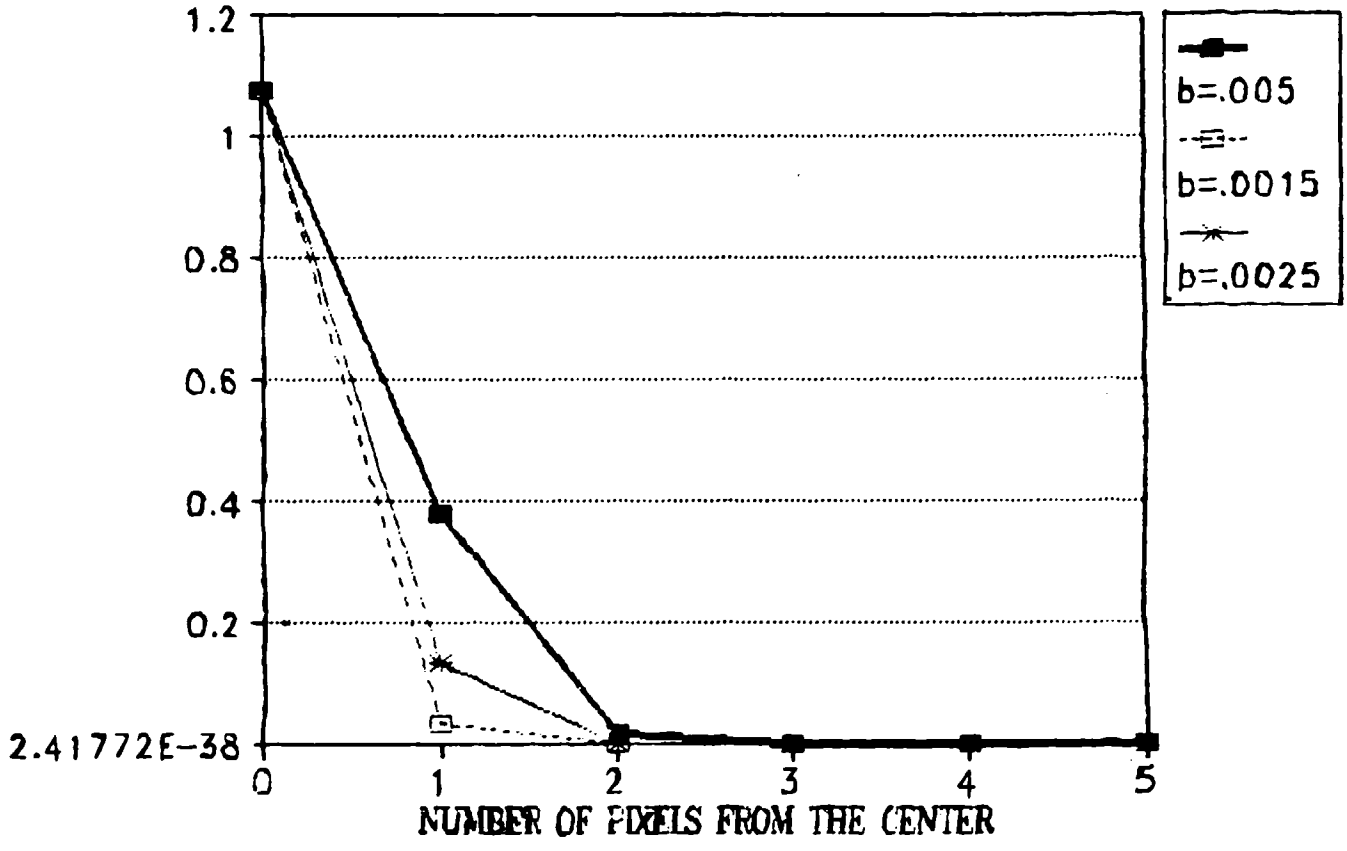


Figure 34
Convolution Filters Corresponding to MTFs in Figure 33

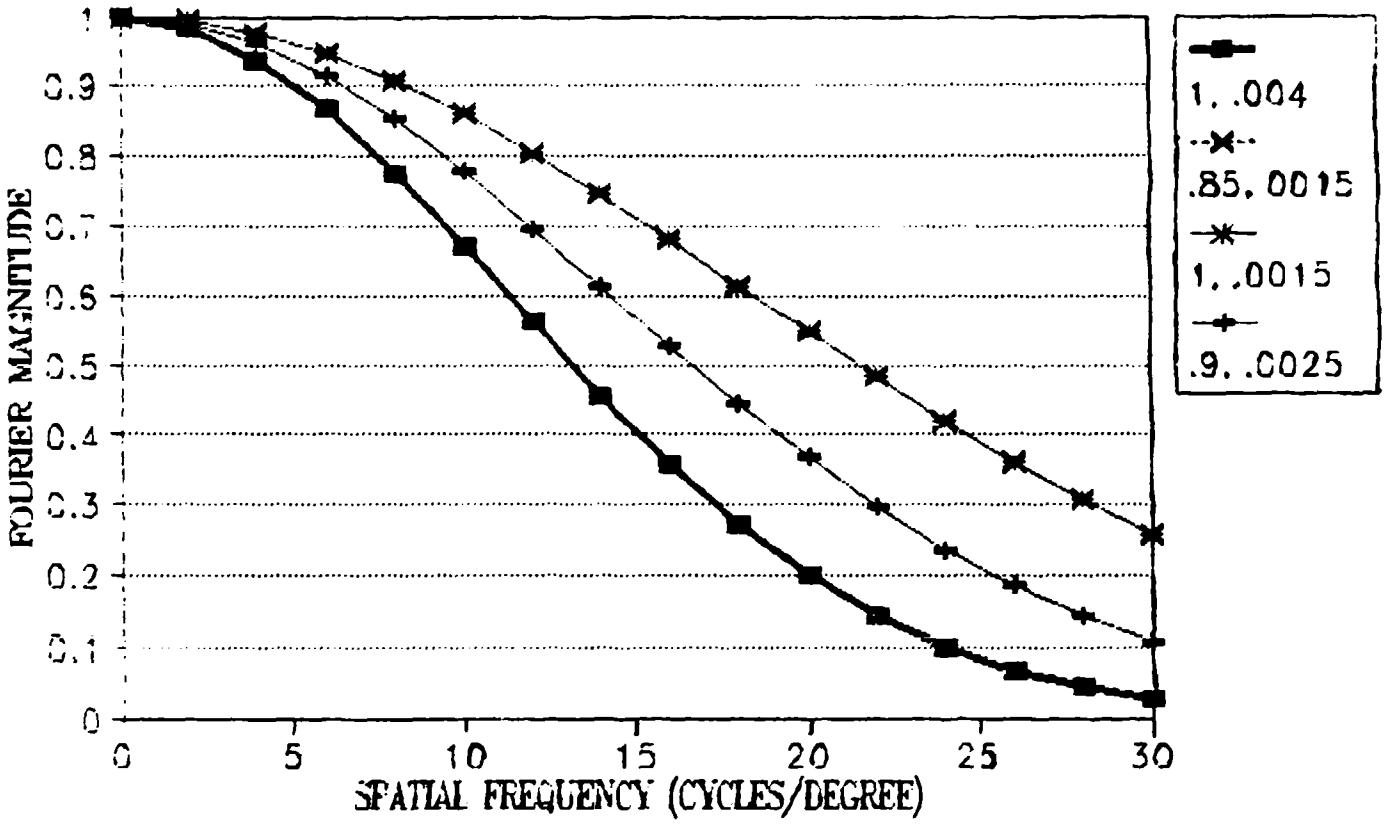


Figure 35
Normalized MTFs for Experimental Filters from Figure 33

where x_1 and x_2 represent distance in the vertical and horizontal dimensions (x is distance in a single dimension) and $h(x_1,x_2)$ is the 2-dimensional convolution filter in the spatial domain ($h(x)$ denotes a unidimensional convolution filter).

The curves in Figure 33 represent hypothetical modulation depth curves which are conceptually interesting. Two of the curves have y-intercept values of 1 (i.e., a = 1) and two curves intercept the y-axis at values lower than 1 (a=.85 and a=.90). Three of the curves cross over one another at approximately 8 cycles/degree of visual angle. Thus, a comparison of these curves would yield an image quality preference for low or high spatial frequency contrast in an image.

As mentioned in the section on measuring luminance from display devices, the filters with y-intercepts of less than 1 in Figure 33 will lower the average luminance of an image. In order to compare MTFs while holding other parameters (e.g., luminance) constant, it becomes necessary to normalize the curves in Figure 33 to a value of 1 at zero frequency. In the spatial domain, this is equivalent to requiring the filter coefficient to sum to a value of 1 (i.e., the area under the curve must integrate to unity). Figure 35 shows the MTF curves in Figure 33 normalized to unity at zero frequency. The result in Figure 35 shows that only three curves from Figure 33 remain distinct and these curves are well ordered in Figure 35. The three distinct MTF curves in Figure 35 may still be used to filter the images in Figures 18 through 22 and the resulting images may be compared for their image quality.

In order to use the filters from Figure 35, the five images shown in Figures 18 through 22 were digitized into 512 by 512 elements. The display device used for the images was a 1,000-line by 1,024-pixel-wide color monitor which was approximately 12 inches in height by 15 inches wide. At a viewing distance of 36 inches, the pixel-to-pixel center distance was approximately dx = .38′ of visual angle (=.023°). With this information and the rightmost equality in Equation 16, a digitized 11 X 11 convolution filter ($h(x_1,x_2)$) was calculated in a radially symmetric fashion with the center element being in the 6th row and 6th column. Coefficients in the convolution filter were solved for by computing their euclidean distance from the center of the filter. For example, the center or highest point in the filter, which shall be denoted as h_{00}, is simply

$$h(0,0) = \frac{a\sqrt{\pi}}{b} e^{-0} = \frac{a\sqrt{\pi}}{b}. \qquad (31)$$

The filter coefficient 4 pixels vertically from the center and 3 pixels horizontally from the center is $h_{4,3}$ and is given by

$$h(4*.023,3*.023) = \frac{a\sqrt{\pi}}{b} e^{-\frac{\pi^2(.092^2+.084^2)}{b}} . \quad (32)$$

For example, if a = 1 and b = .015, $h_{0,0}$ = 118.16 and $h_{3,4}$ = (118.16)(3.67 X 10^{-5}) = .0043 before normalization. In the center of the filter, the exponential part of the equation always evaluates to 1. As we depart from the center of the filter, the exponential part of Equation 30 denotes the contribution of that point in the filter relative to the center of the filter. For example, if b = .015 and we move one pixel horizontally or vertically from the center of the filter, the height of the filter is approximately 70% of the height at the center of the filter. Moving two pixels horizontally or vertically from the center of the filter, the height evaluates to approximately 25% of the center of the filter.

The five 11 X 11 filters representing each of the MTFs was numerically convolved with each of the images. In order to assure no changes in luminance, it was necessary to normalize the matrix of filters such that the sum of the coefficients in the 11 X 11 matrix was 1. Note, as mentioned, that in using this technique, the differential in the DC contrast of the images is reduced to zero, i.e., all DC contrasts are normalized to 1. By restricting the filter weights to sum to 1 in the spatial domain, MTF(0,0) or the heights of the MTFs at DC or zero frequency will automatically evaluate to 1 in the frequency domain. The result is that an MTF crossover effect cannot be simulated unless the images are of different average luminance values.

Displaying the filtered images through a second display device creates the double-pass problem. That is, the images have been filtered to create the effect of interest but the display of these filtered images through another device is a second filtering process. From a linear systems approach, the MTF of the combined processes is the product of the individual MTFs. Therefore, if the original filter MTF is multiplied by the MTF of the display device used, the result is the overall filtering or MTF.

A rough estimate of the display MTF was obtained using the direct measurement method. In the horizontal direction, the limiting mask frequency of the display was 1,024 pixels over approximately 15 inches. Assuming the dark band between each pixel match fills out an on-off cycle, there is a maximum of 1,024 cycles over 15 inches. At a viewing distance of 36 inches, the approximation from Equation 12 states that the horizontal dimension of the display subtends 24 degrees of visual angle. The 1,024 cycles across 24 degrees of visual angle yield approximately 43 cycles/degree of visual angle as a theoretical maximum for

resolution. If the width of the electron beam was constrained to fall inside the holes of the mask, the modulation depth at 43 cycles/degree could be near unity at this limiting spatial frequency. Of course, this type of design (width of electron beam < mask pitch) would most likely make the raster structure of the display device quite visible and distracting (see Murch & Virgin, 1985).

Square waves of varying frequency were displayed on the experimental monitor and modulation at these frequencies was measured using a photometer. Equation 15 was used to estimate the response to a sinusoidal input as a function of the response to a square wave input. Even at square wave frequencies of 5 cycles/ degree, however, the higher order harmonics of Equation 14 are nearly zero. Figure 36 shows the estimated display MTF as well as the double-pass MTFs obtained by multiplying the filters in Figure 35 by the top curve in Figure 36. As can be seen from the resulting MTFs in Figure 36, the double-pass filtering process is limited by the display MTF in Figure 36 relative to the MTFs in Figure 35.

Filtered images were presented side-by-side using a paired comparison approach. At a viewing distance of approximately 30 inches, each image subtended approximately 13 degrees of visual angle in both the horizontal and vertical directions. Note that the viewing distance is already included in the calculation of the filters in Figure 35 as well as the calculation of the display MTF in Figure 36 in order that the MTFs be presented as a function of viewing spatial frequency.

Ambient illumination in the display environment was approximately 1 footcandle and the reflection of this ambient illumination from the display is included in the calculation of the display modulation or MTF in Figure 36.

Casual observation of the paired stimuli revealed that the differences in the images were not detectable at the defined viewing distance (i.e., 36 inches). At much closer viewing distances (denoting a shift in the spatial frequency axis in both Fig. 34 and 35), differences could be detected. In addition, sequential presentation of images directly on top of one another made stimuli discriminable, denoting the importance of the experimental methodology used to present imagery.

In addition, it was evident that filtered images of the Airport scene (see Fig. 18) were more discriminable than filtered versions of the other four images (Fig. 19 through 22). This finding indicates the ineffectiveness of global Fourier analysis (Fig. 23) as a measure of the content of the imagery with respect to viewing.

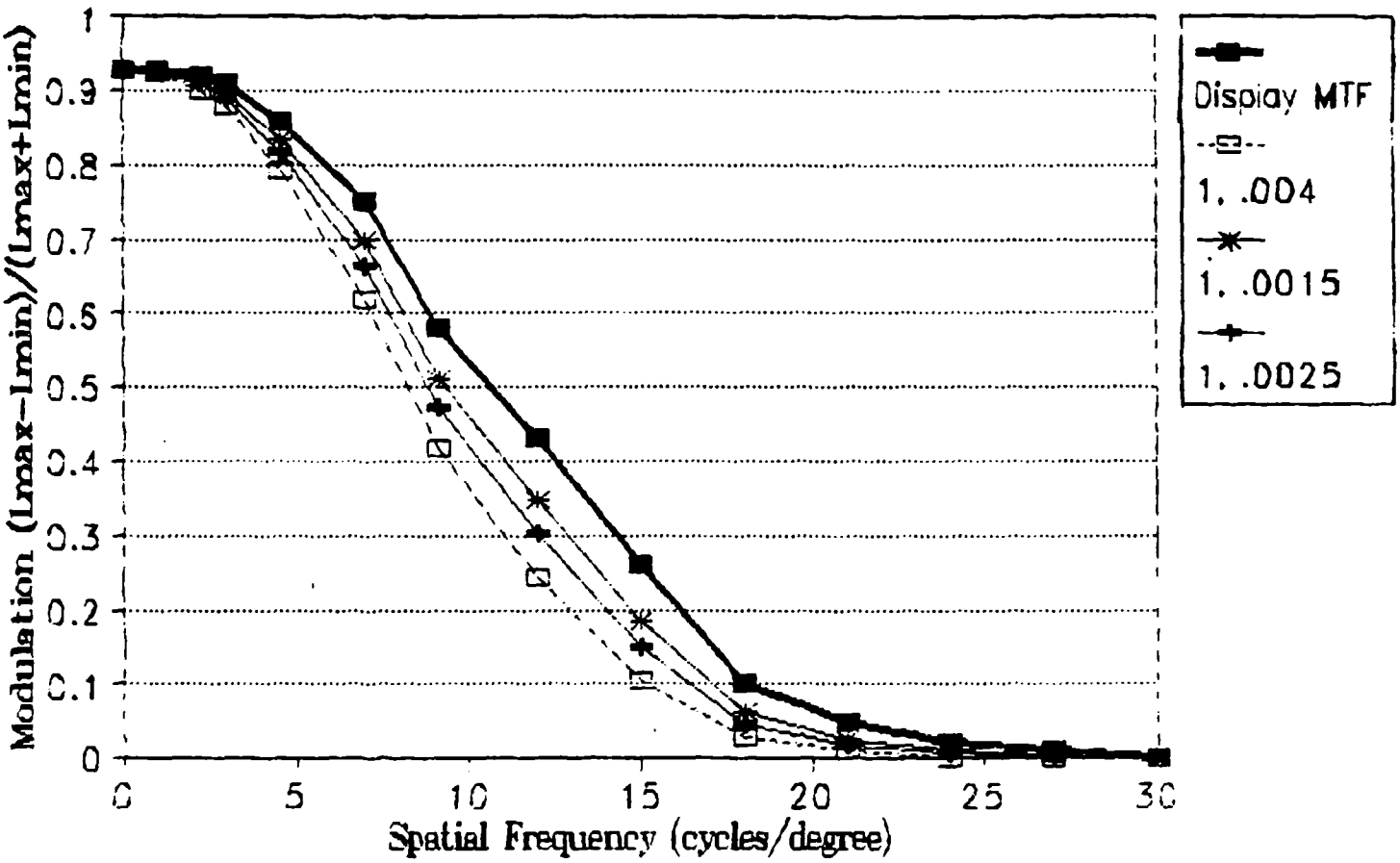


Figure 36
Double-Pass Products of MTFs in Figure 35

The three double-passed MTF curves in Figure 36 differ only by about 3% to 10% modulation between approximately 4 and 18 cycles per degree of visual angle. This range of modulation differences within the mid-frequency range does not serve as a good test of the metrics discussed in this report. The change in the metrics across filters in Figure 36 will typically be less than a single JND. The MTF of the display used in the presentation of the image unexpectedly compressed the filter differences.

DIRECTIONS FOR FUTURE IMAGE QUALITY RESEARCH: IMAGE QUALITY IN A MULTIDIMENSIONAL SPACE

For displays used in visual simulation, image quality preferences are guided by multiple criteria. These criteria reflect the physical variety in available displays which range from helmet-mounted displays to large dome displays. In experimental comparison of these displays, failure to control major factors (e.g., luminance) or eliminate differences while manipulating other factors (display MTF) results in confounded comparisons. The need exists to explore image quality from a multidimensional perspective. However, the capability to manipulate display parameters or factors independently of one another simply does not exist in many experimental situations.

This report has focused on the display MTF as the major driver of image quality. As referenced throughout the report, the MTF has been used as the traditional measure of image quality. In many instances, however, factors such as brightness, field of view, and color appearance may dominate the display MTFs in their contribution to subjective image quality.

As mentioned, it is technically difficult to vary display factors independently of one another. For actual display devices, improvements in one factor typically result in impoverished measures of other factors. For example, larger display areas typically result in lowered luminance, decreased resolution, and impoverished color rendering.

Using the MTF filtering techniques described in this report, a class of display MTFs (i.e., those with luminance modulation near unity at zero spatial frequency) may be simulated for viewing purposes. The filtered image produced by the process represents variation of the MTF factor or dimension. An approach where the MTFs may be systematically varied allows testing of many important hypotheses. For example, as presented in this report, the relative importance of low versus high spatial frequency information (i.e., the weighting of spatial frequency based information) to image quality metrics, is critical to their success.

After systematic manipulation of hypothetical display MTFs, other modifications to the image can then be made. By including variation of the average luminance of the image along with the MTF variation, a factorial comparison may be conducted. Figure 37

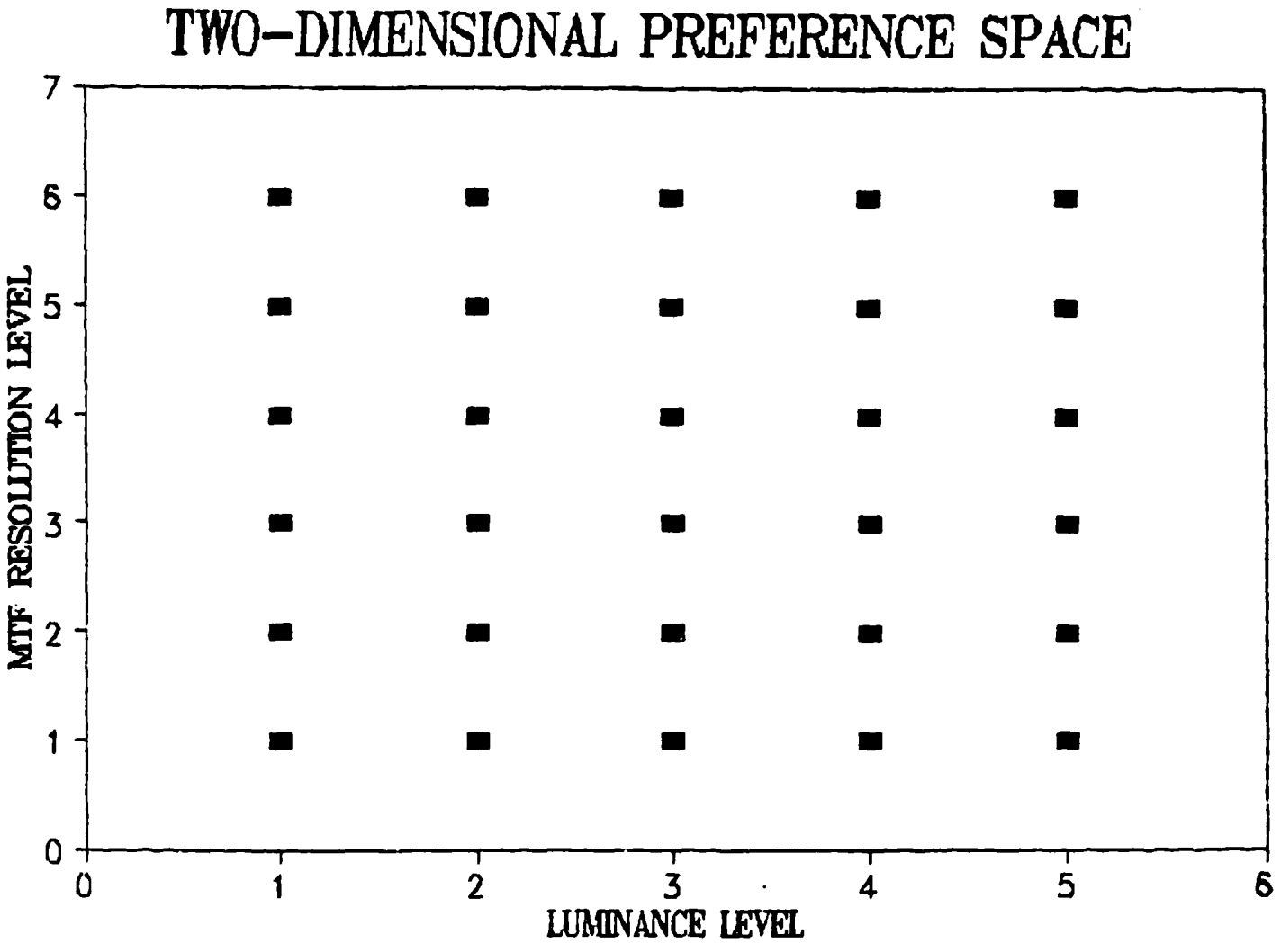


Figure 37
Stimulus Combinations from a 5 X 4
(Display MTF X Luminance) Experiment

represents the stimulus set for a 5 X 4 factorial experiment where display MTF (5 levels) and luminance (4 levels) are varied.

Using a paired comparisons procedure and Likert scale preference ratings (or simply a binary preference) as the response measure, a variety of multidimensional measures may be developed. Figure 38, a two-dimensional isopreference mapping, is one of these measures. Figure 38 denotes a hypothetical mapping where the MTF-luminance combinations in the same region would be equally preferred to each other. If the MTF dimension is permitted to vary in such a fashion as to cover both high- and low-resolution displays, and the luminance dimension is permitted to vary so as to cover bright as well as very dim displays, the mapping in Figure 38 would yield a practical ordinal metric for visual displays covered by the two-dimensional MTF-luminance range.

Luminance and MTF may be varied independently of one another for a constrained subset of pairings (i.e., MTFs which are unity at zero frequency and a variety of luminance and MTF curves which are well within the capability of the display device). Using a more complicated scheme, MTFs may also be generated which need not reach a value of unit at zero spatial frequency. Beyond such manipulations, the next logical parameter of interest is the field of view. Variation of this parameter independently of luminance or MTF is not easily accomplished due to nonhomogeneities introduced for larger field-of-view displays. One possibility is to manipulate the viewing distance and compensate for the shift along the spatial frequency axis (of the image and display content) through software. This procedure would be quite complex, though. In addition, it is not clear that a straightforward experimental comparison can be made across dimensions such as field of view and luminance or MTF. For example, asking an observer whether they prefer a large field-of-view image which is dark as opposed to a small field-of-view image which is bright may simply not be a useful comparison or may not be a meaningful comparison to the observer. With such comparisons, it may also be true that the individual, the task, or the image is a moderating factor in the decision. For example, in sporting bars where viewers can choose large-screen, relatively low-luminance projection devices or small-screen, high-luminance CRTs, the preference may hinge on the type of event being viewed, the event being associated with a task such as tracking the movement of a tennis ball or a football.

Finally, increments in display parameters such as field of view, luminance, and MTF will always be expected or predicted to improve or at least maintain, image quality. However, image-display artifacts (e.g., moire patterns) and display-observer artifacts (e.g., simulator sickness) for specific configurations can contradict these general trends. For example, increasing the field of view of the display at small viewing distances may lower perceived image quality by causing nausea or other simulator sickness-related symptoms. In the near future, these artifacts will most likely be discovered on a case-by-case basis and included in image quality criteria only as exceptions to general models or predictions.

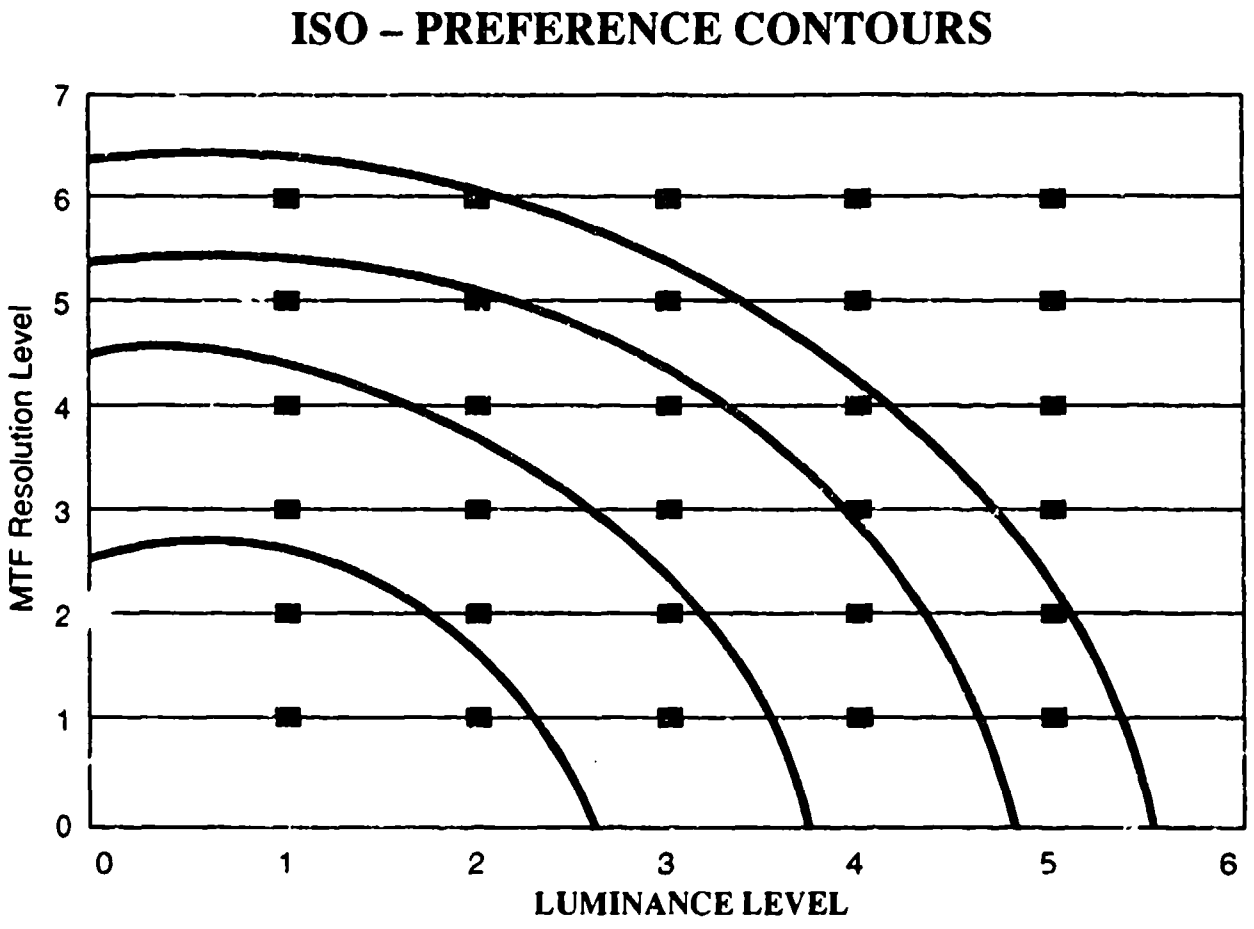


Figure 38
Hypothetical Isopreference Mapping for a
Two-Dimensional Stimulus Space

REFERENCES

Barten, P.G.J. (1984). Spot size and current density distribution of CRTs. Proceedings of the SID, Vol. 25(3), 155-159.

Barten, P.G.J. (1985). Character display capability of CRT monitors. Proceedings of the SID, Vol. 26(4), 293-297.

Barten, P.G.J. (1987). The SQRI method: A new method for the evaluation of visible resolution on a display. Proceedings of the SID, Vol. 28(3), 253-262.

Barten, P.G.J. (1988a) Effects of convergence errors on resolution. Proceedings of the SID, Vol. 29(1), 3-5.

Barten, P.G.J. (1988b) Evaluation of CRT displays with the SQRI method. SID 88 Digest, 445-448.

Barten, P.G.J. (1989). The square root integral (SQRI): A new metric to describe the effect of various display parameters on perceived image quality. SPIE Vol. 1077 Human Vision, Visual Processing, and Digital Display, 73-82.

Barten, P.G.J. (1990). Evaluation of subjective image quality with the square-root integral method. Journal of the Optical Society of America, 7(10), 2024-2031.

Barten, P.G.J. (1991). Evaluation of the effect of noise on subjective image quality. SPIE Vol. 1453 Human Vision, Visual Processing, and Digital Display II, 2-15.

Beaton, R. (1989) Seminar Notes.

Biberman, L. M. (1972). Image Quality. Chapter 2 of Perception of Displayed Information edited by L. M. Biberman, Plenum Press, New York.

Blackwell, H.O. (1946). Contrast thresholds of the human eye. Journal of the Optical Society of America, 36(11), 624-643.

Blakemore, C., Muncey, J., & Ridley, R. M. (1973). Stimulus specificity in the human visual system. Vision Research, 13, 1915-1931.

Campbell, F.W., & Green, D.G. (1965). Optical and retinal factors affecting visual resolution. Journal of Physiology, 181, pp. 376-593.

Campbell, F.W., & Gubisch, R.W. (1966). Optical quality of the human eye. Journal of Physiology, London, 186, 558-578.

Campbell, F.W. & Robson, J.G. (1968). Application of Fourier analysis to the visibility of gratings. Journal of Physiology, London, Vol. 197, 551-566.

Carlson, C.R. (1982). Sine-wave threshold contrast-sensitivity function: Dependence on display size. RCA Review, 43, 675-683.

Carlson, C.R. (1988). Economic display design. Information Display, 4(5), 16-19.

Carlson, C.R., & Cohen, R.W. (1980). A simple psychophysical model for predicting the visibility of displayed information. Proceedings of the SID, 21(3), 229-246.

Evans, R.J. (1990). Image quality metrics and application of the Square Root Integral (SQRI) metric: An overview. (AFHRL-TR-90-56, AD A229 753). Williams Air Force Base, AZ: Operations Training Division, Air Force Human Resources Laboratory.

Field, D.J. (1987). Relations between the statistics of natural images and the response properties of cortical cells. Journal of Optical Society of America, Vol. 4(12), 2379-2394.

Field, D.J. (1989). What the statistics of natural images tell us about visual coding. SPIE Vol. 1077 Human Vision, Visual Processing, and Digital Display, 269-276.

Georgeson, M.A., & Sullivan, G.D. (1975). Contrast constancy: Deblurring in human vision by spatial frequency channels. Journal of Physiology, (London), 252, 627-656.

Glenn, W.E., Glenn, K.G., & Bastian, C.J. (1985). Imaging system design based on psychophysical data. Proceedings of the SID, 26(1), 71-78.

Granger, E.M., & Cupery, K.N. (1972). An optical merit function (SQF), which correlates with subjective image judgments. Photographic Science and Engineering, 16(3), 221-230.

Gubisch, R.W. (1967). Optical performance of the human eye. Journal of the Optical Society of America, Vol. 57, 407-415.

Hood, D.C., & Finkelstein, M.A. (1986). Sensitivity to light. Chapter 5 of Handbook of Perception and Human Performance edited by K.R. Boff, L. Kaufman, and J.P. Thomas, New York: Wiley & Sons.

Hultgren, B.O. (1990). Subjective quality factor revisited. SPIE Vol. 1249 Human Vision and Electronic Imaging: Models, Methods, and Applications, 12-22.

Infante, C. (1985). On the resolution of raster-scanned displays. SID Proceedings, 28(1), 23-35.

Infante, C. (1986). Ultimate resolution and non-Gaussian profiles in CRT displays. Proceedings of the SID, 27(4), 275-280.

Johnson, J. (1958). Analysis of image forming systems. Image Intensifier Symposium, Fort Belvoir, VA, Oct. 6-7, AD 220160. Reprinted in Infrared Design, SPIE Volume 513, Part I edited by R. Barry Johnson and William L. Wolfe, 1985, pp. 761-781.

Kelly, G.R. (1992). Measurement of modulation transfer functions of simulator displays. (AL-TP-1992-0056). Williams Air Force Base, AZ: Aircrew Training Research Division, Armstrong Laboratory.

Kleiss, J.A., & Hubbard, D.C. (1991). Effect of two types of scene detail on detection of altitude change in a flight simulator. (AL-TR-1991-0043, AD A242 034). Williams Air Force Base, AZ: Aircrew Training Research Division, Armstrong Laboratory.

Kulikowski, J.J. (1976). Effective contrast constancy and linearity of contrast sensation. Vision Research, 16, 1419-1431.

Kusaka, H. (January, 1989). Consideration of vision and image quality - Psychological effects induced by picture sharpness. Paper presented at the SPIE/SPSE Conference, Los Angeles, CA.

Murch, G., & Virgin, L. (1985). Resolution and addressability: How much is enough? SID 85 Digest, 101-103.

Peli, E. (1990). Contrast in complex images. Journal of the Optical Society of America, Vol. 7(10), 2032-2040.

Pugh, E.N. Jr. (1988). Vision: Physics and retinal physiology. Chapter 2 of Stevens' Handbook of Experimental Psychology, Volume I: Perception and Motivation, 2nd edition, edited by R.C. Atkinson, R.J. Herrnstein, G. Lindzey, and R.D. Luce, New York: Wiley & Sons., pp. 76-163.

Rogowitz, B.E. (1983). The human visual system: A guide for the display technologist. Proceedings of the SID, 24(3), 235-252.

Roufs, J.A.J. (1989). Brightness contrast and sharpness, interactive factors in perceptual image quality. SPIE Vol. 1077 Human Vision, Visual Processing, and Digital Display, 66-72.

Roufs, J.A.J., & Bouma, H. (1980). Towards linking perception research and image quality. Proceedings of the SID, Vol. 21(3), 247-270.

Schade, O.H., Sr. (1956). Optical and photoelectric analog of the eye. Journal of the Optical Society of America, 46, 721-724.

Schade, O.H. Sr. (1964). An evaluation of photographic image quality and resolving power. Journal of Motion Picture and Television Engineers, 73(2), 81-119.

Schade, O.H., Sr. (June, 1987). Image Quality: A comparison of photographic and television systems. Copyright in 1975 by RCA Corporation and reprinted in Journal of Motion Picture and Television Engineers, 567-595.

Scott, F. (1966). Three-bar target modulation detectability. Photographic Sci. Engineering, 10(1), pp 49-52.

Snyder, A.W., & Srinivasan, M.V. (1979). Human Psychophysics: Functional interpretation for contrast sensitivity versus spatial frequency curve. Biological Cybernetics, 32, 9-17.

Snyder, H.L. (1985). Image quality: Measures and visual performance. Chapter 4 of Flat Panel Displays and CRTs, edited by L.E. Tannas, Jr., Van Nostrand Reinhold Company, Inc., N.Y.

Task, H.L. (1979). An evaluation and comparison of several measures of image quality for television displays. Air Force Aerospace Medical Research Laboratory Technical Report AMRL-TR-79-7.

Thomas, M.L., Reining, G., & Kelly, G. (1990). The display for advanced research and development: An "inexpensive" answer to tactical simulation. SPIE

Vandenberghe, P., De Clercq, A., Schaumont, J., & Bracke, P. (1990). The influence of CRT-Gamma on luminance and modulation. SID 90 Digest, 152-155.

van der Zee, E. & Boesten, M.H.W.A. (1980). The influence of luminance and size on the image quality of complex scenes. IPO Annual Progress Report, Vol. 15, 69-75.

van Meeteren, A. & Vos, J.J. (1972). Resolution and contrast sensitivity at low luminances. Vision Research, Vol. 12, 825-833.

Van Ness, F.L., & Bauman, M.A. (1967). Spatial modulation transfer in the human eye. Journal of the Optical Society of America, 57, 401-406.

Veron, H. (1985). The measurement of resolution of shadow-mask CRTs. Proceedings of the SID, 26(4), 299-303.

Watanabe, A., Mori, T., Nagata, S., & Hiwatashi, K. (1968). Spatial sinewave responses of the human visual system. Vision Research, 8, 1245-1263.

Westheimer, G. (1970). Image quality in the human eye. Optica Acta, Vol. 17(9), 641-658.

Westheimer, G. (1986). The eye as an optical instrument. Chapter 4 of Handbook of Perception and Human Performance edited by K.R. Boff, L. Kaufman, and J.P. Thomas, New York: Wiley & Sons.

Zetsche, C., & Hauske, G. (1989). Multiple channel model for the prediction of subjective image quality. SPIE Vol. 1077 Human Vision, Visual Processing, and Digital Display, 209-216.